



**Etude à l'échelle de la molécule unique des changements  
conformationnels de la molécule d'ADN : influence de la  
présence de défauts locaux présents sur l'ADN et de  
paramètres physico-chimiques de la solution  
environnante**

Annael Brunet

**► To cite this version:**

Annael Brunet. Etude à l'échelle de la molécule unique des changements conformationnels de la molécule d'ADN : influence de la présence de défauts locaux présents sur l'ADN et de paramètres physico-chimiques de la solution environnante. Biophysique. Université Paul Sabatier - Toulouse III, 2015. Français. <NNT : 2015TOU30168>. <tel-01384604>

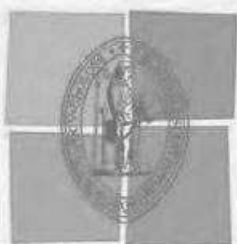
**HAL Id: tel-01384604**

**<https://tel.archives-ouvertes.fr/tel-01384604>**

Submitted on 20 Oct 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université  
de Toulouse

# THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

---

---

Présentée et soutenue le 16/10/2015 par :

**ANNAËL BRUNET**

---

Étude à l'échelle de la molécule unique des changements conformationnels  
de la molécule d'ADN.

*Influence de la présence de défauts locaux présents sur l'ADN et de  
paramètres physico-chimiques de la solution environnante.*

---

---

## JURY

NICOLAS DESTAINVILLE  
CATHERINE TARDIN  
FRANCK JOLIBOIS  
JEAN-FRANCOIS ALLEMAND  
ANNICK LESNE  
CENDRINE MOSKALENKO  
AURÉLIEN BANCAUD  
MANOEL MANGHI

Pr, LPT, Toulouse  
MC, IPBS, Toulouse  
Pr, LPCNO, Toulouse  
Pr, LPS-ENS, Paris  
DR, LPTMC, Paris  
MC, LP-ENS, Lyon  
CR, LAAS-CNRS, Toulouse  
MC, LPT, Toulouse

Directeur de thèse  
Directeur de thèse  
Président du Jury  
Rapporteur  
Rapporteur  
Examinatrice  
Invité  
Invité

---

École doctorale et spécialité :

*Sciences De la Matière : Physique et Bio-Physique - COR 02*

Unité de Recherche :

*Laboratoire de Physique Théorique (LPT-IRSAMC, UMR UPS/CNRS 5152)*

*Institut Pharmacologie et de Biologie Structurale (IPBS, UMR CNRS/UPS 5089)*

Directeur(s) de Thèse :

*Nicolas DESTAINVILLE et Catherine TARDIN*

Rapporteurs :

*Jean-Francois ALLEMAND et Annick LESNE*



UNIVERSITÉ PAUL SABATIER, TOULOUSE III  
LPT ET IPBS

# THÈSE

présentée en première version en vue d'obtenir le grade de Docteur,  
spécialité « Physique et Bio-physique »

par

Brunet Annaël

## ÉTUDE À L'ÉCHELLE DE LA MOLÉCULE UNIQUE DES CHANGEMENTS CONFORMATIONNELS DE LA MOLÉCULE D'ADN.

*Influence de la présence de défauts locaux  
présents sur l'ADN et de paramètres  
physico-chimiques de la solution environnante.*

Thèse soutenue le 16/10/2015 devant le jury composé de :

M.	NICOLAS DESTAINVILLE	LPT	(Directeur)
M <sup>me</sup>	CATHERINE TARDIN	IPBS	(Directrice)
M.	FRANCK JOLIBOIS	LPCNO	(Examinateur)
M.	JEAN-FRANCOIS ALLEMAND	LPS-ENS	(Rapporteur)
M <sup>me</sup>	ANNICK LESNE	LPTMC	(Rapportrice)
M <sup>me</sup>	CENDRINE MOSKALENKO	LP-ENS	(Examinatrice)
M.	AURÉLIEN BANCAUD	LAAS-CNRS	(Invité)
M.	MANOEL MANGHI	LPT	(Invité)



*À qui vous voulez...*





# REMERCIEMENTS

**S**I, aujourd'hui, je revenais 3 ans et 2 mois en arrière en gardant la mémoire physique et mentale de cette thèse, je crois bien que je re-signerai sans hésiter ! Et si l'ensemble de ces années est synonyme pour moi d'une formidable aventure tant scientifique qu'humaine, cela ne fut possible que grâce à de nombreuses personnes. Avec un grand respect, une intemporelle gratitude et une profonde affection, j'espère réussir à vous remercier, autant que je le souhaite, par ces quelques mots.

Tout d'abord, il fut nécessaire d'être acceptée officiellement au sein de mes deux laboratoires toulousains, qui ont quelques fois fait office de résidence secondaire, pour cela je remercie donc Jean-Philippe Girard directeur de l'IPBS ainsi que Bertrand Georgeot directeur de l'IRSAMC. Je tiens aussi à remercier Clément Sire qui a toujours été un si convivial et dynamique directeur du LPT, je n'oublierai jamais cet appel aux : "SEMINAIRE !".

Merci à toi, Catherine, de m'avoir appris cette rigueur et cette justesse scientifique, de toujours m'avoir m'inclus dans les discussions et réflexions en lien avec ce projet mais aussi d'autres domaines de recherche. Tu n'as jamais hésité à me faire part de ce que tu pensais de mon travail et je t'en remercie car si aujourd'hui je me suis affirmée sur ce projet, c'est en très grande partie grâce à toi et ta confiance. Je tiens aussi à insister, en le posant sur papier, que toujours, vraiment toujours, tu as été disponible sur le plan professionnel et personnel.

Merci à Nicolas d'avoir toujours été si attentif et si patient, surtout au début, pour me réexpliquer plusieurs fois les mêmes choses car, il faut avouer que vos discussions avec Manoel vont souvent vite pour une personne tierce ! Merci à Manoel pour ton exigence mathématique et d'avoir toujours été si bienveillant. Vous avez toujours été incroyablement accessibles.

Même si je trouve que suis encore trop loin du niveau scientifique de vous tous, mes chefs, j'espère y tendre un jour et l'ensemble des questions que vous m'avez posé durant ma thèse, pour me pousser à avancer toujours plus loin, ainsi que vos conseils avisés, m'y aideront.

Merci à l'ensemble des membres de mon jury : Franck Jolibois, Jean-Francois Allemand, Annick Lesne, Cendrine Moskalenko, Aurélien Bancaud et Manoel Manghi d'avoir accepté de juger mon travail, écrit comme oral, de vos questions et remarques constructives ainsi que d'avoir pris le temps, à la suite de la soutenance, de me délivrer quelques précieux conseils pour mon futur scientifique. J'espère un jour vous annoncer que je soutiens ma HDR et qu'elle fait moins que 400 pages !

Merci à Laurence d'avoir fait de l'équipe à l'IPBS une véritable famille et m'avoir dit depuis longtemps, même s'il m'a fallu du temps pour réellement m'en rendre compte, que je menais ce projet avec adresse. Merci à tous les autres membres de l'équipe, permanents ou non, à Sébastien et Cheikh de m'avoir formé au début de ma thèse, Serge pour ton aide et toutes ces madeleines que tu m'aménais durant mes journées de manip, Evert pour tous ces livres témoins de ta curiosité, Fabrice pour ta bonne humeur habituelle même en parlant de rugby, Etienne pour tes croques, Pascalou pour tout, Thomas, Meredith, Nelson et Mathilde pour votre si grande sympathie. Merci aussi à tous les autres membres de l'équipe du LPT.

Merci à Aurélien Bancaud d'avoir, dès le premier jour où je t'ai annoncé que je postulais à cette thèse qui m'intéressait tellement, téléphoné à Catherine et Nicolas pour soutenir ma candidature et ceci a conduit, heureusement pour moi, à modifier leurs plans. Merci à Marie Brut, Anne Heymerick et Georges Landa, rencontrés à cette époque et dont je souhaite (et dont je sais) que je resterai toujours en contact.

Mais, soyons honnête, cette histoire a commencé bien plus tôt...

Merci à Mr Jourdain de m'avoir montré, les cheveux au vent, que la physique "c'est fun quand même !" ; à Mme Beguet de m'avoir dit, peut-être pas exactement en ces termes mais pas très loin, "si tu veux faire une thèse, fais la !" ; à Nicolas Suaux de m'avoir donné un aperçu de la recherche (en L2 !) sans cela je n'aurais jamais continué dans cette voie ; à Sébastien Dutour de m'avoir rappelé, quand je doutais, que nous sommes acteurs de notre histoire. Merci Erick Grouteau de m'avoir offert une liberté.

Merci à toute ma famille, tout ce que je pourrais dire ne sera jamais assez par rapport à tout ce que je voudrais vous dire, tout ce que vous avez fait pour moi et surtout tout ce que vous représentez pour moi. À Aline pour tous tes choix qui ont permis de me construire et de m'ouvrir un champ infini des possibles, de m'avoir toujours soutenu, bravant les moulins à vent quelque soit la fatigue pour que j'aille jusqu'au bout de mes choix. À Aurélie d'être le samouraï que tu es, inconditionnellement prête à tout pour moi. À Jean-Partick d'avoir, maladroitement parfois, mais au fond toujours, timidement, été là. À Aymeric d'avoir discrètement été là, toujours, plus que je ne l'ai été. À Mamette de toujours faire sensation, surtout en public, en ayant plus la tête dans les étoiles que moi. Je ne serais pas une personne aussi forte qui peut toujours garder le sourire sans l'un(e)s d'entre vous.

Merci à Bertil, Mijo, Guillaume, Cécile, Mamy et Yannick même si nous ne nous voyons pas si souvent, vous étiez là. Merci à tout le reste de la famille et des cousins (on est cousins au deuxième ou troisième degré déjà ? je ne sais plus !). Merci à tous les autres aussi, Aude, Anne, Anne, Audrey, Laurent, Carole...

A word cloud of names arranged in a circular shape, representing the names of people who have been in the French national football team. The names are of various sizes and orientations, creating a dense, circular pattern. Some names are repeated, such as 'Nanou', 'Julie', and 'Pascalou'.

Merci à tous ceux que j'ai oublié, je m'en excuse platement mais quelque part cet oubli m'aide à ne pas faire 10 pages de remerciements !

Le plus sincèrement du monde, je vous remercie, sûrement les larmes aux yeux, de tout ce que vous m'avez aidé à construire.

Toulouse, le 25 novembre 2015.

vii



# RÉSUMÉ DE LA THÈSE

TITRE : ÉTUDE À L'ÉCHELLE DE LA MOLÉCULE UNIQUE DES CHANGEMENTS CONFORMATIONNELS DE LA MOLÉCULE D'ADN.

*Influence de la présence de défauts locaux présents sur l'ADN et de paramètres physico-chimiques de la solution environnante*

Les ions jouent un rôle majeur sur les processus biologiques affectant la molécule d'ADN que ce soit en termes d'activité de liaison de protéines à l'ADN ou d'encapsulation de l'ADN dans les capsides virales ou le noyau. L'activité de protéines sur l'ADN est, par ailleurs, fréquemment liée à une courbure locale de l'axe de la double hélice, que ce soit en raison d'une séquence intrinsèquement courbée, ou, via la capacité de protéines à courber la séquence sur laquelle elles se fixent. Être capable de caractériser et comprendre l'effet des ions présents en solution, de la courbure et de la dénaturation locale de la molécule d'ADN sur les conformations de cette dernière est donc crucial pour approfondir la compréhension de nombreux processus biologiques. Des travaux, tant expérimentaux que théoriques, ont déjà été menés sur ces questions mais celles-ci sont encore largement débattues. En effet, pour y répondre, doivent notamment être développées des méthodes expérimentales qui ne perturbent pas significativement la conformation de l'ADN ou le complexe ADN-protéine, ainsi que des modèles théoriques associés permettant une analyse précise des données expérimentales et leur compréhension physique.

L'objectif de ce travail est de proposer des outils expérimentaux et théoriques permettant de décrire physiquement l'influence de défauts locaux présents sur la molécule d'ADN et de paramètres physico-chimiques de la solution environnante.

A cette fin, des données expérimentales ont été acquises à l'échelle de la molécule unique grâce à la technique haut-débit de "Tethered Particle Motion" (HT-TPM). Le TPM consiste à enregistrer, au cours du temps, les positions d'une particule accrochée à l'extrémité d'une molécule d'ADN, immobilisée par son autre extrémité sur un support en verre. L'utilisation d'une biopuce permettant la parallélisation des complexes ADN/particule et l'acquisition "à haut débit" de données TPM a permis d'obtenir une grande accumulation de statistiques individuelles. Une procédure d'analyse efficace a été élaborée afin de déterminer les amplitudes du mouvement des assemblages ADN-particules valides. En parallèle, ont été effectuées des simulations basées sur un modèle de physique statistique mésoscopique dans lequel la molécule d'ADN est assimilée à une chaîne de billes de rayons variables dont les déplacements sont régis par la diffusion brownienne et une énergie potentielle d'interaction prenant en compte notamment l'énergie de courbure du polymère ADN.

Une première étude a porté sur l'effet de la force ionique de la solution environnante sur la longueur de persistance ( $L_p$ ), qui traduit la rigidité du polymère d'ADN. Les valeurs de  $L_p$  extraites des données de HT-TPM ont fait apparaître une décroissance de la longueur de persistance de 55 à 30 nm, corrélée à l'augmentation de la force ionique, avec une décroissance plus forte observée pour les ions divalents  $Mg^{2+}$  que pour les ions monovalents  $Na^+$ . Les valeurs de  $L_p$  déterminées sur une plage étendue de force ionique ont permis de valider l'approche théorique proposée par Manning en 2006 dans la cas  $Na^+$ .

Une deuxième étude a conduit à l'élaboration d'une méthode permettant de quantifier l'angle de courbure locale induite par une séquence spécifique ou la liaison d'une protéine sur la molécule d'ADN. L'échantillon modèle a été obtenu en insérant de une à sept séquences CAAAAACGG en phase. Une description théorique de la chaîne d'ADN appelée "kinked Worm-Like Chain" a été proposée. Elle conduit à une formule simple de la distance bout-à-bout de l'ADN qui permet d'extraire la valeur de l'angle de courbure à partir des mesures de HT-TPM. Ainsi, il a pu être montré que la séquence CAAAAACGG induit un angle de  $19^\circ \pm 4^\circ$  en accord avec les données de la littérature.

Une troisième étude concernant la mesure de l'impact de la dénaturation partielle de l'ADN, induite par la température, sur sa rigidité apparente globale a été menée. Des résultats préliminaires sont proposés.

TITLE : STUDY AT THE SINGLE MOLECULE LEVEL OF CONFORMATIONAL CHANGES OF THE DNA MOLECULE.

*Impact of local defects included in the DNA molecule and of a large set of physicochemical conditions*

Ions play an important role in many biological processes affecting the DNA molecule, both for binding activities of DNA-protein interaction, and the DNA packaging in viral capsids or in the cell nucleus. Proteins actions on DNA are also often associated to the double helix curvature, be it because of an intrinsic curved sequence, or of the ability of the proteins, to curve the sequence they are trying to bind. Being able to characterize and understand the effects on the DNA conformation of ions present in solution, DNA local curvature, and local denaturation bubble is essential and crucial for the thorough understanding of many biological processes. Many experimental, and theoretical studies have already been conducted to address these questions. However they remain highly debated. To answer then one must notably develop experimental approaches that minimize alteration of the conformation of the DNA molecule or the complex protein-DNA, as well as associated theoretical models that permit a precise analysis of experimental data as well as their physical understanding.

The goal of this work is to develop and propose experimental and theoretical tools which would provide a physical description of the influence of DNA local defects on the DNA molecule as well as of physicochemical conditions of the DNA environmental solution.

For this purpose, experimental data have been collected, at a single molecule level, using the High-Throughput Tethered Particle Motion" (HT-TPM) technique. TPM consists of recording the location of a particle grafted by one end of a single DNA molecule and immobilized, at the other end, to a glass surface. The use of a biochip that enables the parallelization of DNA/particle complexes and the ensuing high-throughput data acquisition permitted to obtain a large accumulation of individual statistics . A strong analysis procedure has been developed to extract and quantify the amplitude of motion of the valid DNA/particle complexes . Alongside that, simulations have been run, based on a mesoscopic statistical mechanics model in which the DNA molecule is related to a chain of monomers with varying radius and in which the amplitude of motion is governed by both the Brownian motion and by the interaction potential associated to stretching and bending energies of the polymer.

A first study was conducted on the effect of the ionic strength induced by surrounding ions in solution on the DNA persistence length ( $L_p$ ) which characterizes the DNA polymer rigidity. The extracted  $L_p$  values of HT-TPM measurement decrease from 55 to 30 nm when the ionic strength increases. A stronger decrease was observed in presence of divalent ions  $Mg^{2+}$  than with monovalent ions  $Na^+$ . This quantification of  $L_p$  dependence, on a large and strongly prospected range of ionic strengthes, tends to validate the theoretical approach proposed in 2006 by Manning in presence of monovalent ions  $Na^+$ .

A second project allows us to develop a method of evaluation and quantification of local DNA bending angles, induced either by specific intrinsic sequence, or by the binding of proteins on DNA. Constructs made of 575 base-pair DNAs with in-phase assemblies of one to seven sequences CAAAAAACGG was used. A theoretical description of the polymer chain, named "kinked Worm-Like Chain" was proposed which leads to a simple formulation of the end-to-end distance of DNA molecules allowing to extract local bend angles from HT-TPM measurement. As a result, we find that the sequence CAAAAAACGG induces a bend angle of  $19^\circ \pm 4^\circ$ , in agreement with other value from the literature.

A third study concerning the influence of temperature-induced partial denaturation on the global apparent rigidity parameters of the polymer was conducted. Preliminary results are proposed.



# TABLE DES MATIÈRES

<b>Remerciements</b>	<b>v</b>
<b>Résumé de la thèse</b>	<b>ix</b>
RÉSUMÉ . . . . .	ix
ABSTRACT . . . . .	xi
<b>TABLE DES MATIÈRES</b>	<b>xiii</b>
<b>GLOSSAIRE</b>	<b>xvii</b>
ABRÉVIATIONS . . . . .	xvii
NOTATIONS . . . . .	xvii
CONSTANTES . . . . .	xviii
RÈGLES DE CONVERSION . . . . .	xviii
<b>Introduction générale</b>	<b>1</b>
<b>I L'ADN : de la molécule au polymère</b>	<b>5</b>
1 LA MOLÉCULE D'ADN	7
INTRODUCTION . . . . .	8
1.1 IN VIVO . . . . .	9
1.2 L'ACIDE DÉSOXYRIBO NUCLÉIQUE . . . . .	11
1.3 LA DOUBLE HÉLICE . . . . .	19
1.4 LA NOTION DE FLEXIBILITÉ DE L'ADN DANS LA CELLULE . . . . .	28
CONCLUSION . . . . .	31
2 LE POLYMÈRE ADN	33
INTRODUCTION . . . . .	34
2.1 MODÈLES STATISTIQUES DE POLYMÈRE . . . . .	35
2.2 MODÈLES FLEXIBLES . . . . .	35
2.3 MODÈLES SEMI-FLEXIBLES . . . . .	42
2.4 CHAÎNE RÉELLE . . . . .	46
CONCLUSION . . . . .	49
<b>Conclusion</b>	<b>51</b>
<b>Bibliographie</b>	<b>53</b>
BIBLIOGRAPHIE . . . . .	53
<b>II Notre approche à l'échelle de la molécule unique d'ADN</b>	<b>57</b>
3 TECHNIQUE EXPÉRIMENTALE : <i>Tethered Particle Motion (TPM)</i>	59

INTRODUCTION . . . . .	61
3.1 PRINCIPE DU <i>Tethered Particle Motion (TPM)</i> . . . . .	63
3.2 TPM-ON-A-CHIP . . . . .	68
3.3 POINTS FORTS DU TPM-ON-A-CHIP . . . . .	69
3.4 DISPOSITIF EXPÉRIMENTAL . . . . .	70
3.5 PRINCIPE D'UNE MESURE TPM . . . . .	76
3.6 FILTRES DE SÉLECTION APPLIQUÉS PAR LE LOGICIEL <i>NanoMultiplex</i> . . . . .	81
3.7 CORRECTIONS ET PREMIÈRES ÉTAPES DE L'ANALYSE DE DONNÉES . . . . .	85
3.8 MESURES TPM . . . . .	100
CONCLUSION . . . . .	104
 4 MODÉLISATION DE L'ADN . . . . .	 105
INTRODUCTION . . . . .	107
4.1 SIMULATION MONTE-CARLO DYNAMIQUE (MC-D) . . . . .	109
4.2 SIMULATION PAR ÉCHANTILLONNAGE EXACT (EE) . . . . .	125
CONCLUSION . . . . .	131
 Objectifs de la thèse . . . . .	 133
 Bibliographie . . . . .	 135
BIBLIOGRAPHIE . . . . .	135
 III Effet Global : variation de la force ionique . . . . .	 139
 5 ÉTUDE EXPÉRIMENTALE ET THÉORIQUE DES VARIATIONS DE LA FORCE IONIQUE SUR L'ADN . . . . .	 141
INTRODUCTION . . . . .	143
5.1 LES POLYÉLECTROLYTES . . . . .	145
5.2 MESURE DE LA VARIATION DES PROPRIÉTÉS PHYSIQUES DE MOLÉCULES D'ADN EN FONCTION DE LA FORCE IONIQUE . . . . .	154
5.3 VARIATION DE $L_p$ . . . . .	163
5.4 VALEUR DE LA LONGUEUR DE PERSISTANCE PRÈS DES CONDITIONS SA- LINES PHYSIOLOGIQUES . . . . .	163
5.5 THÉORIE ODIJK, SKOLNICK, FIXMAN . . . . .	165
5.6 PONDÉRATION À FAIBLE $I$ DE LA THÉORIE OSF PAR LE PARAMÈTRE DE MANNING . . . . .	166
5.7 CORRECTION AUX VALEURS DE $I$ INTERMÉDIAIRES PAR UNE THÉORIE DE CHAMP MOYEN NON LINÉAIRE . . . . .	170
5.8 PRISE EN COMPTE DES FORCES D'ÉTIREMENT DANS UNE NOUVELLE THÉORIE DE MANNING . . . . .	172
5.9 PROPOSITION D'UNE FORMULE D'INTERPOLATION . . . . .	175
5.10 MESURE D'UN EFFET SÉQUENCE PAR L'APPROCHE TPM . . . . .	177
5.11 OBSERVATION SUR LA VARIATION DU TEMPS DE RELAXATION . . . . .	182
CONCLUSION . . . . .	183
ARTICLE . . . . .	185
 Bibliographie . . . . .	 203
BIBLIOGRAPHIE . . . . .	203

<b>IV</b>	<b>Effet Local : courbure et bulle de dénaturation</b>	<b>205</b>
6	COURBURE LOCALE INTRINSÈQUE À L'ADN	207
	INTRODUCTION . . . . .	209
6.1	SÉQUENCES A-TRACT, INDUISANT UNE COURBURE INTRINSÈQUE DANS LA MOLÉCULE D'ADN . . . . .	211
6.2	THÉORIE POUR UNE COURBURE LOCALE . . . . .	212
6.3	MESURE DE LA VARIATION DE LA DISTANCE BOUT-À-BOUT DES COMPLEXES ADN/PARTICULE EN FONCTION DE LA SÉQUENCE . . . . .	215
6.4	EXTRACTION DE LA DISTANCE BOUT-À-BOUT DE L'ADN SEUL . . . . .	215
6.5	QUANTIFICATION DE L'ANGLE DE COURBURE À PARTIR D'UN JEU DE CONSTRUCTIONS D'ADN . . . . .	217
6.6	PROPOSITION D'UNE FORMULE SIMPLE AFIN D'ÉVALUER L'ANGLE DE COURBURE À PARTIR DE DEUX MESURES . . . . .	220
6.7	THÉORIE POUR UN INSERT RIGIDE LOCAL . . . . .	222
	CONCLUSION . . . . .	226
	ARTICLE . . . . .	229
7	DÉTECTION D'UNE BULLE DE DÉNATURATION EN TPM	253
	INTRODUCTION . . . . .	255
7.1	L'ADN, UN BIO-POLYMÈRE SENSIBLE À LA TEMPÉRATURE . . . . .	257
7.2	MODÈLES DE DÉNATURATION . . . . .	261
7.3	STRATÉGIES EXPÉRIMENTALES . . . . .	267
7.4	PROCÉDURE EXPÉRIMENTALE . . . . .	273
7.5	INFLUENCE DE LA TEMPÉRATURE SUR L'ÉTAPE DE CORRECTION DU MOYENNAGE VIDÉO . . . . .	276
7.6	INFLUENCE DE LA TEMPÉRATURE SUR LES PROPRIÉTÉS PHYSIQUES D'UNE MOLÉCULE D'ADN DE 2060 BP . . . . .	282
7.7	COMPARAISON AVEC DES DONNÉES TPM DE LA LITTÉRATURE . . . . .	286
	CONCLUSION . . . . .	293
	<b>Bibliographie</b>	<b>295</b>
	BIBLIOGRAPHIE . . . . .	295
	<b>CONCLUSION GÉNÉRALE &amp; PERSPECTIVES</b>	<b>297</b>
<b>V</b>	<b>Annexes</b>	<b>309</b>
A	PROTOCOLE . . . . .	313
B	ADNs UTILISÉS . . . . .	324
C	SCRIPT <i>Mathematica</i> D'ANALYSE DES DONNÉES . . . . .	331
D	SCRIPT DES PROGRAMMES NUMÉRIQUES . . . . .	364
E	DONNÉES GÉNÉRALES . . . . .	366
F	DONNÉES SUR L'EFFET DE LA FORCE IONIQUE . . . . .	367
G	DONNÉES SUR L'EFFET D'UNE COURBURE LOCALE . . . . .	375
H	DONNÉES SUR L'EFFET DE LA TEMPÉRATURE . . . . .	375



# GLOSSAIRE

## ABRÉVIATIONS

ADN	Acide DésoxyriboNucléique	ARN	Acide RiboNucléique
A	Adénine	T	Thymine
C	Cytosine	G	Guanine
bp	paire de base		
ss	ADN simple brin		
ds	ADN double brins		
$C_i$	Carbone $i$	$O_i$	Oxygène $i$
BSA	Albumine de Sérum Bovin	$N_i$	Azote $i$
PEG	PolyÉthylène Glycol		
PDMS	PolyDiMéthylSiloxane		
FJC	Freely Jointed Chain		
FRC	Freely Rotating Chain		
KR	Modèle de Kratky-Porod		
WLC	Worm Like Chain		

## NOTATIONS

$pK_a$	Mesure de la constante d'activité d'un acide
$N$	Nombre de monomères, ou segments, constituant la chaîne d'ADN
$l$	Longueur des monomères, ou segments, constituant la chaîne d'ADN
$\mathbf{r}_i$	Vecteur définissant l'orientation du monomère, ou segment, $i$
$L_{ADN}$	Longueur de la molécule d'ADN en bp, assimilable à $L$
$L = Nl = N_{bp} \times 0.34$	Longueur de contour de la molécule d'ADN en nm, $L = L_{ADN} \times 0.34$
$P(\mathbf{R}, N)$	Probabilité d'arriver au point $\mathbf{R}$ en partant de $(0, 0, 0)$ en $N$ pas.
$\langle \mathbf{R}^2 \rangle$	Écart quadratique moyen entre les deux extrémités de la chaîne
$R = \sqrt{\langle \mathbf{R}^2 \rangle}$	Distance bout-à-bout de la chaîne
$\langle \mathbf{R}_g^2 \rangle$	Écart quadratique moyen entre la position d'un monomère et le centre de masse $G$ de la chaîne
$R_g = \sqrt{\langle \mathbf{R}_g^2 \rangle}$	Rayon de giration de la chaîne
$S$	Entropie de la chaîne
$b$	Longueur statistique de Kuhn
$L_p$	Longueur de persistance
$\kappa_c$	Module élastique de courbure
$D$	Coefficient de diffusion
$\zeta$	Coefficient de friction
$\eta$	Viscosité du milieu
$\nu$	Exposant de Flory

$\tau_{  }$	Temps de relaxation des complexes ADN/particule
$R_{TPM} = \sqrt{\langle \mathbf{R}_{TPM}^2 \rangle}$	Distance bout-à-bout des complexes ADN/particule mesurée par TPM
$R_{Exp  } = \sqrt{\langle \mathbf{R}_{Exp  }^2 \rangle}$	Distance bout-à-bout des complexes ADN/particule corrigée de l'effet de moyennage vidéo, dépendant de $\tau_{  }$
$R_{ADN}$	Distance bout-à-bout de l'ADN, assimilable à $R$

## CONSTANTES

$k_B$	$1.38110 \times 10^{-23} \text{ J.K}^{-1}$	Constante de Boltzmann
$E_{them} = k_B T$	$4.1 \times 10^{-21} \text{ J}$	Énergie d'agitation thermique à $25\text{C}^\circ$
$N_{Genome}$	$3 \times 10^9 \text{ bp}$	Nombre de paire de base de notre génome
$\eta_{eau}$	$1.15 \times 10^{-3} \text{ Pa.s}$	Viscosité de l'eau
$N_a$	$6.022 \times 10^{23} \text{ mol}^{-1}$	Nombre d'Avogadro
$e$	$1.602 \times 10^{-19} \text{ C}$	Charge électrique élémentaire d'un proton
$\epsilon = \epsilon_0 \times \epsilon_1$	$\text{C}^2\text{N}^{-1}\text{m}^{-2}$	Permittivité du milieu
$\epsilon_0$	$8.85 \times 10^{-12} \text{ C}^2\text{N}^{-1}\text{m}^{-2}$	Permittivité du vide
$\epsilon_1$	78.4	Permittivité relative pour l'eau à $25\text{C}^\circ$
$R_{DNA}$	$1 \times 10^{-9} \text{ m}$	Rayon d'un double brin d'ADN de forme <b>B</b>
$A_{DNA}$	$0.17 \times 10^{-9} \text{ m}$	Densité linéique de charge. Ici, distance moyenne entre deux charges élémentaires le long d'un double brin d'ADN.

## RÈGLES DE CONVERSION

1 bp	$\Rightarrow$	0.34 nm
1 kcal/mol	$\Rightarrow$	$6.95 \times 10^{-21} \text{ J}$
1 $k_B T$	$\Rightarrow$	$4.1 \times 10^{-21} \text{ J}$

# INTRODUCTION GÉNÉRALE

LA molécule d'ADN, communément représentée et connue sous la forme d'une double hélice, est un bio-polymère possédant une grande variabilité structurale qui influe sur de nombreux processus biologiques comme les phénomènes d'expression et de régulation des gènes, ou encore lors de la compaction de l'ADN dans la noyau cellulaire chez les eucaryotes. Certains d'entre eux peuvent être associés à une courbure locale intervenant dans la molécule d'ADN ou à l'ouverture locale de la double hélice qui forme ainsi une bulle de dénaturation. Ces fluctuations structurales peuvent être liées à l'activité de protéines sur l'ADN, le plus souvent associées à la formation d'une courbure locale et ce, que ce soit en raison d'une séquence intrinsèquement courbée favorable à la fixation de protéine, ou via la capacité de protéines à courber la séquence sur laquelle elles sont venues se fixer. Le phénomène de dénaturation peut être provoqué par différents procédés, notamment par des agents chimiques ou physiques comme les protéines, ou causé par un accroissement de la température. Nous nous intéressons ici à la formation de bulles de dénaturation en fonction de la température. De plus, les ions jouent un rôle majeur sur les processus biologiques affectant la molécule d'ADN que ce soit, là encore, en termes d'activité de liaison de protéines à l'ADN ou bien lors de l'encapsulation de l'ADN dans le noyau ou dans les capsides virales.

L'objectif de ce travail est donc de proposer des outils expérimentaux et théoriques permettant de décrire physiquement l'influence de défauts locaux de la molécule d'ADN et de paramètres physico-chimiques de la solution environnante.

Dans ce travail, nous proposons de combiner approches théoriques et expérimentales. Pour cela, nous avons tiré parti de nos expériences et simulations numériques à l'échelle de la molécule unique et exploité les différents modèles physiques existants pour décrire les polymères. Nous avons montré qu'il était possible d'accéder à la mesure d'un angle de courbure locale, ou à la variation du paramètre caractérisant la rigidité de la molécule d'ADN, sa longueur de persistance, sur des molécules d'ADN présentant des défauts locaux ou soumises à diverses conditions physico-chimiques.

La *première partie* présente notre objet d'intérêt qu'est le bio-polymère d'ADN. Pour pleinement comprendre et appréhender dans leur ensemble les règles et contraintes associées à sa structure et à sa description il est nécessaire de partir de l'objet biologique avant d'arriver au système physique. À cet égard, le premier chapitre de cette première partie est consacré à la description de la molécule biologique d'ADN. Les principales contraintes structurales sont explicitées afin de nous permettre de comprendre la notion de flexibilité et rigidité structurale inhérente à cet objet. Ensuite, le deuxième chapitre détaille les différents modèles physiques qui nous permettent de caractériser et appréhender la physique et la mécanique régissant les différentes propriétés des polymères.

La *deuxième partie* est consacrée à nos approches expérimentales et numériques à l'échelle de la molécule unique et nous permet de définir pleinement la problématique de ce projet de thèse. Dans un premier chapitre l'ensemble de la méthode expérimentale nommée "Thetered Particle Motion" (TPM) de suivi de molécule unique est présenté, ainsi que l'apport de notre bio-puce nommée "TPM-on-a-chip", qui nous permet de suivre en parallèle et en temps réel environ 500 trajectoires associées à nos molécules d'ADN. Cette accumulation de données est un formidable atout afin d'explorer l'ensemble des conformations statistiques de la molécule d'ADN avec une forte précision expérimentale. De plus, l'ensemble des étapes de la procédure d'analyse que nous avons mise en place durant cette thèse y est proposé. Le deuxième chapitre de cette partie explicite la physique et les outils physiques utilisés afin de modéliser le polymère d'ADN en fonction de la géométrie et des contraintes expérimentales. Deux modèles numériques ont été développés. Le premier est un modèle de physique statistique mésoscopique, basé sur un algorithme de Monte Carlo, qui nous permet d'explorer, dans certaines conditions favorables, la dynamique associées à nos systèmes. Le deuxième modèle dit d'échantillonnage exact ne rend plus compte de la dynamique mais nous permet d'explorer entièrement l'espace des configurations accessibles à notre système. Ce dernier présente un gain de temps indéniable pour sonder des phénomènes purement statistiques. La description précise qui est proposée dans ces deux chapitres, nous permet de soupeser les points forts et les limites de ces deux approches, et nous permet de mettre en lumière la complémentarité et l'apport scientifique de la combinaison de ces dernières, et de l'ensemble des approches dont nous disposons. L'ensemble de ces deux modèles nous permettent de résoudre le problème inverse pour extraire les propriétés physiques et mécaniques associées à notre molécule d'ADN à partir des données expérimentales.

Dans les troisième et quatrième parties, sont présentés les résultats obtenus sur les trois projets, menés de front, au cours de cette thèse.

- Le premier consiste à sonder un effet global induit par des changements des conditions physico-chimiques de la solution environnant l'ADN. Plus précisément, ce projet correspond à l'étude de l'effet de la force ionique sur la longueur de persistance du bio-polymère en présence d'ions monovalents  $Na^+$  et/ou d'ions divalents  $Mg^{2+}$ .
- Le deuxième projet consiste à explorer l'effet local induit, soit par la présence d'une séquence spécifique, ici la séquence CAAAAACGG, provoquant une courbure locale et ponctuelle intrinsèque, soit par la fixation d'une protéine sur la molécule d'ADN.
- Enfin, le dernier projet prospecte l'influence de la température et la formation de bulles de dénaturation sur la rigidité de la molécule d'ADN. Cet effet est la combinaison d'un effet global suivi, à plus haute température, de l'apparition de défauts locaux lors de l'ouverture de bulles de dénaturation dans la séquence. Ce dernier projet est constitué de résultats encore préliminaires.



À l'heure actuelle, cette thèse a fait l'objet de deux publications sur les deux premiers projets que nous venons de décrire.

- **Brunet, A.**, Tardin, C., Salomé, L., Rousseau, P., Destainville, N., Manghi, M., *Dependence of DNA persistence length on ionic strength of solutions with monovalent and divalent salts : a joint theory-experiment study*, **Macromolecules**, 2015, 48 (11), pp 3641–3652, DOI : 10.1021/acs.macromol.5b00735
- **Brunet, A.**, Chevalier, S., Destainville, N., Manghi, M., Rousseau, P., Salhi, M., Salomé, L., Tardin, C., *Probing a label-free local bend in DNA by single molecule tethered particle motion*, **Nucleic acids research**, 2015, 42 (11), pe 72(7) DOI : 10.1093/nar/gkv201



Première partie

**L'ADN : de la molécule au  
polymère**



# LA MOLÉCULE D'ADN

# 1

## SOMMAIRE

INTRODUCTION . . . . .	8
1.1 IN VIVO . . . . .	9
1.1.1 L'information génétique . . . . .	9
Transcription . . . . .	9
Réplication . . . . .	10
1.2 L'ACIDE DÉSOXYRIBO NUCLÉIQUE . . . . .	11
1.2.1 Vision générale . . . . .	11
1.2.2 Structure des nucléotides . . . . .	12
Le phosphate . . . . .	12
Le sucre . . . . .	12
Bases azotées . . . . .	14
Le groupement phosphodiester . . . . .	15
Résumé des corrélations conformationnelles . . . . .	16
Orientation d'un simple brin . . . . .	17
1.2.3 Structures des bases complémentaires . . . . .	17
Appariement de Chargaff . . . . .	17
Grand et petit sillons . . . . .	19
1.3 LA DOUBLE HÉLICE . . . . .	19
1.3.1 Pourquoi une forme en double hélice ? . . . . .	19
1.3.2 Structures d'empilement des bases complémentaires . . . . .	20
Empilement des bases complémentaires . . . . .	20
Modulation angulaire autour des bases complémentaires . . . . .	21
Orientation de la double hélice . . . . .	21
1.3.3 Les formes principales de double hélice . . . . .	22
Forme canonique B . . . . .	23
Forme canonique A . . . . .	24
Forme Z . . . . .	25
Principales variations des paramètres structuraux . . . . .	27
1.4 LA NOTION DE FLEXIBILITÉ DE L'ADN DANS LA CELLULE . . . . .	28
1.4.1 La compaction de l'ADN . . . . .	28
1.4.2 La notion de flexibilité . . . . .	29
CONCLUSION . . . . .	31

## INTRODUCTION

**A**cide Désoxyribo Nucléique (ADN), c'est le nom d'un ingrédient clé de la vie sur Terre. En effet, la molécule d'ADN porte l'information génétique autrement dit le **génome** [Ave44]. Cette macromolécule est l'élément central et fondamental autour duquel s'articulent tous les processus liés à l'activité cellulaire. Des *instructions* sont inscrites et codées sur la molécule d'ADN, ces dernières organisent précisément les structures et les processus mis en jeu dans la matière biologique. Cela comprend **l'expression et la régulation des gènes**. Cette molécule d'ADN contient les gènes, qui possèdent chacun une séquence, une position nommée *locus* le long de la molécule, et codent une fonctionnalité bien spécifique. Par sa capacité de se répliquer, l'ADN permet la transmission de ces informations qui constituent le matériel héréditaire.

L'ADN code toutes les informations nécessaires à la synthèse et à l'arrangement hiérarchique multi-échelles, qu'elles soient **spatiales** ou **temporelles**, de toutes les cellules qui nous composent. Lors des processus intervenant *in vivo*, la molécule d'ADN peut interagir avec d'autres biomolécules, modifiant ou déformant ainsi sa structure locale. Cela implique une possibilité de **modulations locales**, spatiales et/ou temporelles, de sa structure. Autrement dit, la molécule d'ADN possède une certaine **flexibilité structurale** lui permettant cette réponse mécanique. Quelles sont les caractéristiques qui confèrent, à cette dernière, de telles propriétés physiques et mécaniques ? Quels paramètres structuraux lui confèrent cette flexibilité ?

La compréhension et l'enchaînement de ces mécanismes biologiques passent par la description de la structure et de la topologie de la molécule d'ADN. Dans ce chapitre nous commencerons par présenter l'organisation atomique, puis moléculaire de l'ADN, afin de décrire sa structure primaire. Nous distinguerons très vite l'existence d'un panel de conformations accessibles. Cette richesse de conformations découle de la structure même de l'ADN et induit une **flexibilité structurale** de l'objet. Nous enchaînerons ensuite, sur son arrangement tridimensionnel. Ceci nous permettra de mettre en lumière les limites ainsi que les contraintes influant sur cette **flexibilité structurale**. Enfin, nous présenterons les principales règles qui régissent les interactions avec d'autres molécules. Ainsi nous commencerons à toucher du doigt l'importance même de la séquence de l'ADN qui, nous le verrons plus tard, peut impacter son comportement global.

## 1.1 IN VIVO

L'ADN est une formidable banque de données. L'évidence que l'ADN porte l'information génétique dans la cellule fut initialement prouvée par O. T. Avery en 1944 [Ave44], puis confirmée par d'autres études [Her52, MS58]. Le *dogme central* de la biologie moléculaire définit l'ADN, l'ARN<sup>1</sup> et les protéines comme les trois macromolécules cruciales dans la *diffusion* de cette information génétique.

### 1.1.1 L'information génétique

Lors de cette phase de *diffusion*, les *instructions* sont transmises à de nouveaux acteurs cellulaires : l'ARN et les protéines. Ce flux d'informations est unidirectionnel et est défini selon les étapes-clés suivantes, Fig. 1.1 :

- la **transcription** : transfert de l'information génétique entre l'ADN à l'ARNm<sup>2</sup>
- la **traduction** : transfert de l'information génétique de l'ARNm aux protéines.

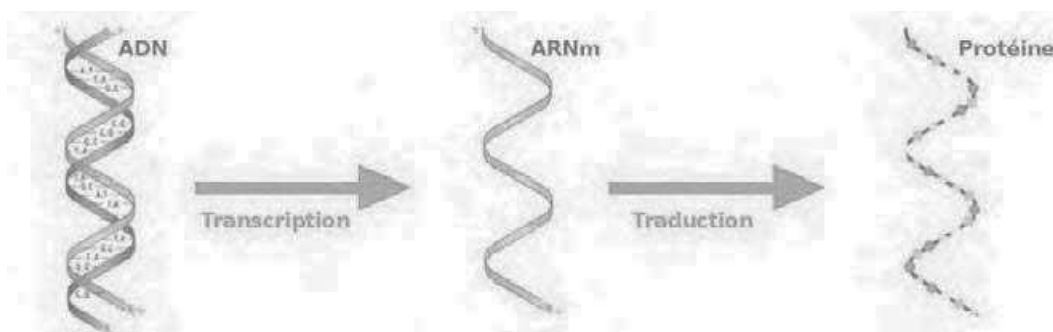


Fig. 1.1 – Selon le dogme de la biologie moléculaire, schéma de l'expression des gènes.

De nombreux opérateurs protéiques ou enzymatiques interviennent lors de chacune de ces étapes.

### Transcription

Le phénomène de transcription est possible grâce à une ouverture de la structure en double-hélice de l'ADN, le détail de celui-ci apparaît dans la suite de ce manuscrit. Cette déstabilisation s'opère notamment dans les régions riches en bases azotées Adénine et Thymine. En effet, comme nous le verrons par la suite, de part la nature de ces bases azotées il est plus facile de rompre leurs interactions. La séparation des brins de l'ADN favorise la fixation et l'insertion de complexes enzymatiques grâce au gain en *volume libre* proche des brins d'ADN. Ainsi, ce phénomène de transcription modifie localement, à l'échelle spatiale et temporelle, l'arrangement de la molécule d'ADN.

1. Acide Ribo Nucléique (ARN), molécule simple brin

2. Acide Ribo Nucléique messager (ARNm), molécule simple brin copie d'une partie de l'ADN qui est ensuite traduite en protéines.

Nous nous intéressons aux processus intervenants uniquement sur notre objet d'intérêt qu'est l'ADN, nous laisserons donc de côté le processus de traduction. L'expression du programme génétique est réalisée lors de la phase de **transcription**. Intéressons-nous maintenant à la transmission de l'information héréditaire, processus biologique correspondant à la **réplication**.

### Réplication

Comme son nom l'indique, lors de cette autre phase biologique l'ADN est dupliqué en sa parfaite copie<sup>3</sup>. Cette réplication de l'ADN intervient lors de chaque division cellulaire.

Là encore, ce processus n'est possible que via une ouverture de la structure de l'ADN. S'ajoute à cela l'apparition d'une contrainte de torsion dans la molécule, liée à la réplication même. Ce phénomène fait intervenir un grand nombre de protéines qui se fixent et interagissent avec l'ADN. Ces protéines interviennent, soit lors de la phase de *lecture* de la séquence, soit pour favoriser et maintenir l'avancement de cette lecture. Cet ensemble de phénomènes entraîne, là encore, des modifications locales et temporaires de l'agencement de la double-hélice [FANF11]. En effet, sa structure en amont et en aval de la réplication n'est pas exactement la même, sans compter l'importante modification qui s'opère au cœur même du nœud de réplication.

Ces grands processus liés à la vie cellulaire, débutent tous deux par une ouverture locale de la structure. De plus, la position même de cette première modification n'est pas anodine. Tout ceci souligne l'importance du détail de la structure de l'ADN. En effet, de nombreuses études montrent le rôle **spécifique** de plusieurs **séquences locales** et courtes dans l'ADN [Hag90, SRR76]. L'impact de la séquence locale de l'ADN est primordial car certaines séquences permettent l'initiation de la transcription et/ou de la réplication, ou permettent la fixation de protéines. Et, cela laisse entrevoir l'intérêt et le rôle de la séquence locale sur l'ensemble de la molécule d'ADN.

Suite au décodage de l'information génétique, découlent la synthèse, l'assemblage et la régulation de chaque constituant biologique. Tout ceci souligne bien la place clé qu'a l'information génétique et donc qu'a l'ADN dans tous les phénomènes biologiques.

La description et la compréhension des propriétés physiques et mécaniques de la molécule d'ADN sont donc cruciales pour pleinement comprendre les mécanismes intervenants dans le déroulement et le fonctionnement des processus biologiques, ainsi que leurs enchaînements et progression. De plus, tout ceci soulève l'influence de la signature d'un effet local sur le comportement global de l'ADN.

3. Des erreurs sont possibles lors de la réplication d'ADN tel qu'un appariement incorrect. Ce taux d'erreur de 1 pour 10000 bases répliquées, diminue grâce au mécanisme continu de réparation s'opérant sur la molécule d'ADN.



## 1.2 L'ACIDE DÉSOXYRIBO NUCLÉIQUE

Les informations contenues sur la molécule d'ADN codent pour la construction, le développement et le fonctionnement des biomolécules formant les matériaux élémentaires du vivant, notamment l'ARN et les protéines. La flexibilité structurale, existant au sein de cette molécule, dépend de ces constituants. Intéressons-nous aux différentes caractéristiques de chacun des constituants qui confèrent à cette dernière de telles propriétés.

La molécule d'ADN, structure restée longtemps inconnue, fut l'objet de longues investigations. Celles-ci aboutirent en 1953 [WC53] à la fameuse structure en double-hélice proposée par les scientifiques **J. D. Watson** et **F. Crick**. Le détail de celle-ci repose sur l'exploitation de clichés de diffraction par rayons X, réalisés par R. Franklin et R.G. Gosling [FG53b], et ne fut possible que grâce à la contribution de nombreux autres acteurs scientifiques. Ainsi, pour pleinement comprendre les détails de la structure de la molécule d'ADN, il est nécessaire d'avancer pas à pas en précisant les contributions des scientifiques tels que E. Chargaff [Cha50], W.T. Astbury [Ast39] Rosalind Franklin et Raymond G. Gosling [FG53b, FG53a], ou encore Maurice Wilkins [WSW53].

### 1.2.1 Vision générale

Partons de la vue générale de l'ADN, qu'un grand nombre d'entre-nous partageons. L'ADN est composé de deux **simples** brins interagissant fortement l'un avec l'autre. Chacun de ces simples brins est constitué d'une succession d'unités nommées **nucléotides**, Fig. 1.2 ①. L'**appariement** et l'enchevêtrement de ces deux brins de polymères s'opèrent via des **liaisons hydrogènes** et forment ainsi une structure régulière en forme de **double hélice**, Fig. 1.2 ②.

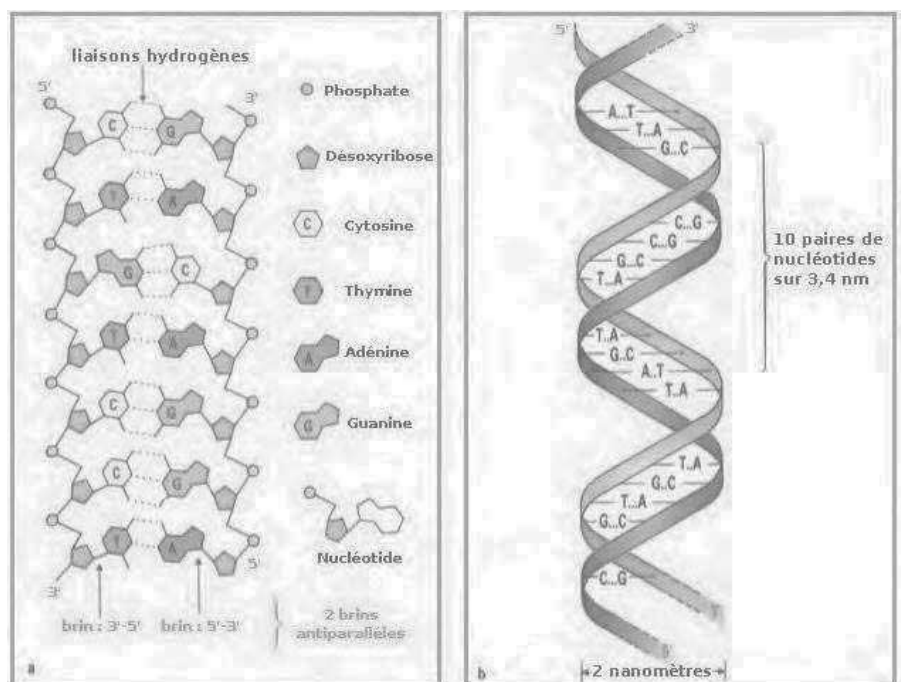


Fig. 1.2 – Représentation d'une partie de la molécule l'ADN [Dio89]. ① Agencement de chaque brin constituant l'ADN, seulement si celui-ci est supposé déroulé. ② Sur cette représentation chaque brin s'enroule autour de l'autre, formant ainsi une structure en double hélice

La molécule d'ADN est une macromolécule complexe. Seule, la discussion de ses particularités peut nous permettre de comprendre les forces qui régissent sa structure. Cette double hélice est-elle la structure d'équilibre ? Quelles sont les limites de sa stabilité ? Ou encore, quels paramètres extérieurs peuvent influencer sur sa structure ?

### 1.2.2 Structure des nucléotides

Chaque simple brin d'ADN est constitué d'un enchaînement de *motifs élémentaires* nommées **nucléotides**, Fig. 1.3. Un nucléotide est composé d'un **acide phosphorique**, d'un **sucré désoxyribose** et d'une **base azotée**. Les **nucléotides** composent le **squelette** de la molécule d'ADN.

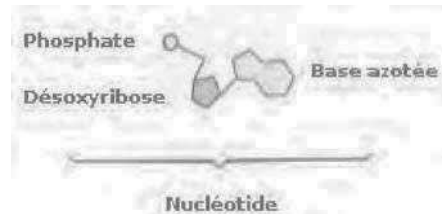


Fig. 1.3 – Schéma d'un nucléotide [Dio89], en accord avec Fig. 1.2 (a)

#### Le phosphate

Une molécule de phosphate se lie au sucre via une liaison phospho-ester O-C. En solution, un phosphate porte une charge élémentaire négative, due à l'ionisation d'un groupement OH, Fig. 1.4. Ainsi chaque phosphate constituant un nucléotide laisse accessible un électron, ce qui confère un caractère négatif à la molécule d'ADN.

Insistons dès maintenant sur ce point : la molécule d'ADN est chargée négativement. Une chaîne simple brin d'ADN possède un électron toutes les bases, ce qui implique qu'une chaîne double brins présente deux électrons par paire de bases.

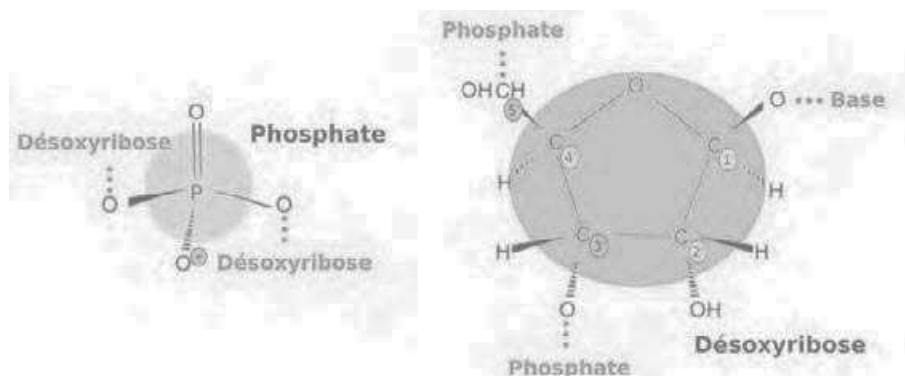


Fig. 1.4 – Représentation, en solution, d'une molécule de phosphate et d'une molécule de sucre désoxyribose en interaction avec les autres constituants de l'ADN. La charge négative portée par le phosphate ainsi que les carbones 5' et 3' du sucre, définissant l'orientation d'un brin d'ADN, sont mis en évidence.

#### Le sucre

Le sucre des acides nucléiques est constitué d'un cycle de cinq atomes. Contrairement à l'image statique que cela pourrait suggérer, celle-ci n'est pas parfaitement plane. En effet, le sucre est comme *plissé*, dû à l'existence de deux types de déformations, schématisées en Fig. 1.5 :

- Forme *Enveloppe E*, un atome est déplacé.
- Forme *Twist T*, deux atomes sont déplacés.

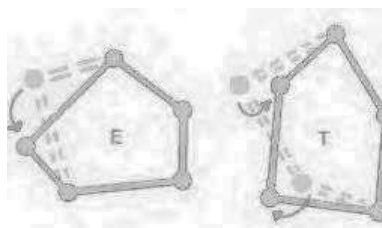
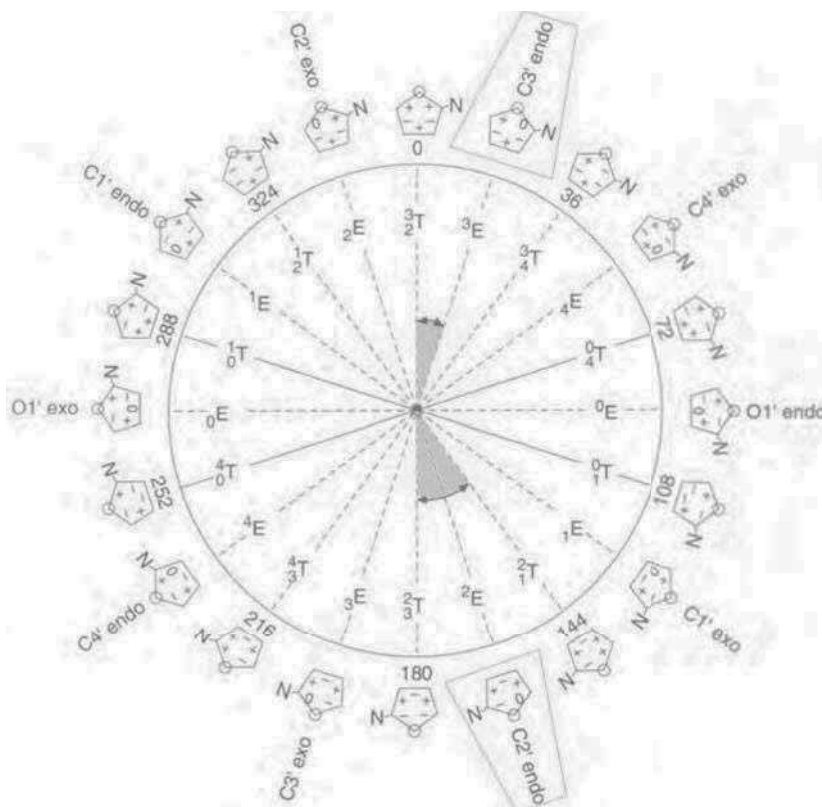


Fig. 1.5 – Déformations possibles du désoxyribose [Dau93]

Chaque déformation existe soit par déplacement de l'atome *vers le haut*, déformation *endo*, ou *vers le bas*, déformation *exo*, et ce pour chacun des cinq atomes constituant le cycle. Il existe donc dix formes *E* et dix formes *T*, car une déformation *Twist* déplace simultanément les deux atomes.

L'ensemble de ces déformations, décrivant les *plissements* du cycle de la molécule de désoxyribose, est représenté sur le diagramme de *pseudorotation* [AS72] ci-dessous Fig. 1.6. Celui-ci représente les 20 conformations possibles. Elles sont disposées sur un cercle de  $360^\circ$ , séparées de façon équidistante suivant un angle de phase de  $18^\circ$ .

Le détail de ce diagramme de *pseudorotation* n'est pas le point clé de ce discours. Ici, ce paragraphe a pour but de mettre en exergue la diversité des conformations accessibles pour un seul des constituants d'un nucléotide. De plus, un nucléotide est répété autant de fois que le nombre de bases qui constitue la molécule d'ADN d'intérêt.

Fig. 1.6 – Cercle de *pseudorotation* recensant les structures accessibles pour la molécule de désoxyribose proposée en 1972 par C. Altona et M. Sundaralingam [AS72], extrait de [BCT00]

Nous ne décrivons pas l'ensemble des conformations accessibles. Nous noterons simplement que les valeurs les plus fréquentes de l'angle de phase de la molécule de sucre se situent entre  $0^\circ$  et  $18^\circ$ , structure nommée **C3'-endo**, et entre  $144$  et  $180$  degrés, structure nommée **C2'-endo**. En solution, les molécules solubles de sucre sont en équilibre dynamique entre ces deux conformations principales.

Des estimations de la variation énergétique du désoxyribose en fonction des conformations présentes sur le cycle de pseudorotations ont été réalisées [MA78].

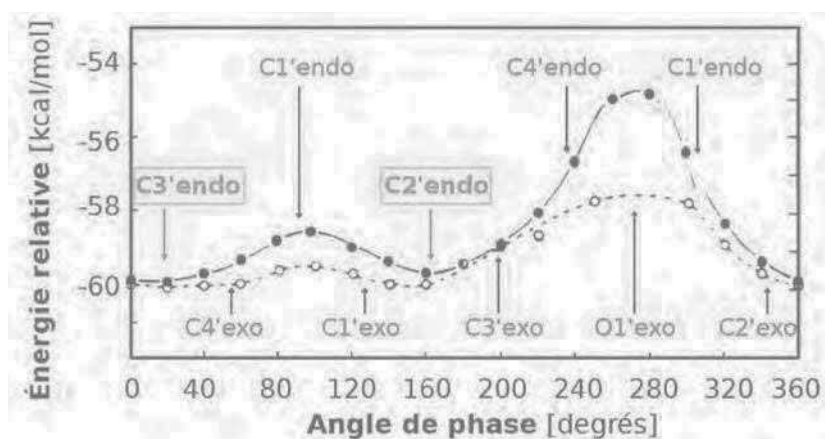


Fig. 1.7 – Variation de l'énergie totale de la molécule de sucre en fonction de l'angle de phase [MA78]. Les deux courbes se différencient par les contraintes angulaires appliquées. La courbe solide prend en compte les corrélations entre tous les angles de torsion de la molécule de sucre. Tandis que la courbe pointillée correspond à un système étudié sans corrélation angulaire.

Via une méthode de champ moyen, les conformations les plus favorables, conformations de plus basses énergies, correspondent bien aux conformations **C3'-endo** et **C2'-endo**. En considérant qu'il n'existe pas de corrélations entre les axes de rotation de la molécule de sucre, la barrière énergétique pour passer de l'état **C3'-endo** à **C2'-endo** est de  $0.5 \text{ kcal.mol}^{-1} \Rightarrow 3.5 \times 10^{-21} \text{ J}$ , courbe pointillée et marqueurs vides en Fig. 1.7. Cette valeur est de l'ordre de l'énergie d'agitation thermique à  $25^\circ$ ,  $E_{\text{them.}} = k_B T_{298K} = 4.1 \times 10^{-21} \text{ J}$ , où  $k_B$  est la constante de Boltzmann. Si les corrélations des angles de rotation internes de la molécule sont ajoutées, courbe pleine et marqueurs pleins en Fig. 1.7, la barrière énergétique passe alors à  $\sim 1.5 \text{ kcal.mol}^{-1}$  soit  $\sim 2.5.E_{\text{them.}}$ .

### Bases azotées

Les bases azotées interagissent avec le sucre désoxyribose par une liaison N-osidique, C-N. Ces dernières sont au nombre de quatre : Adénine notée **A**, Thymine notée **T**, Guanine notée **G** et Cytosine notée **C**.

Il existe deux grandes familles de nucléotides. Les **bases pyrimidiques**, telles que la cytosine ou la thymine, possèdent une structure dérivant de la molécule de pyrimidine, cela signifie qu'elles sont constituées d'un seul noyau aromatique. Tandis que les **bases puriques**, telles que l'adénine et la guanine, dérivant de la structure de la purine, sont constituées de deux cycles aromatiques. Ci-dessous, Fig. 1.9 et 1.8, les structures chimiques de ces bases azotées sont représentées.

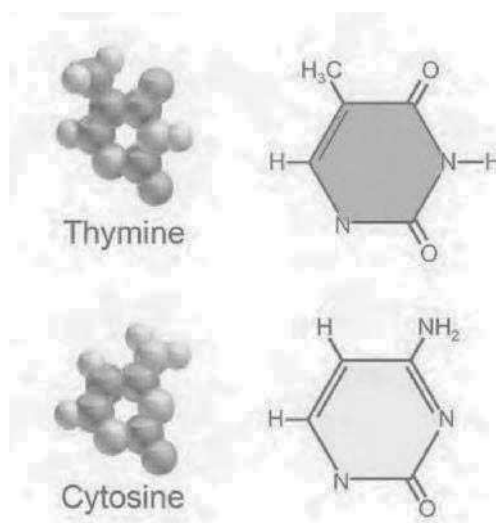


Fig. 1.8 – Représentation des Pyrimidines

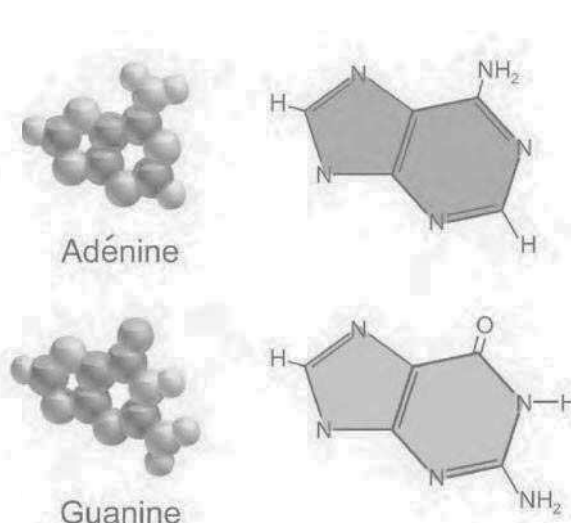
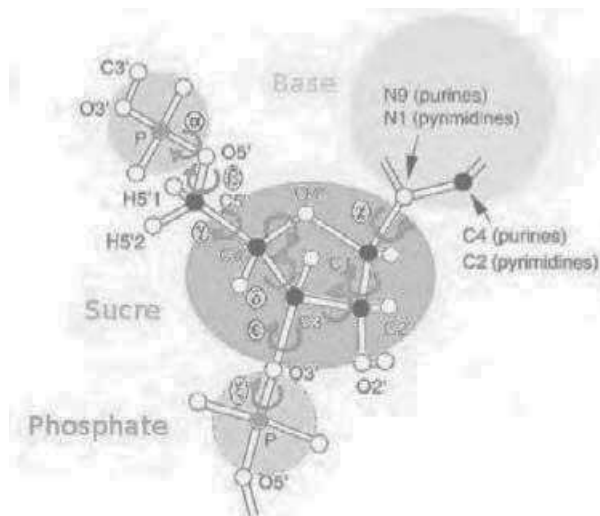


Fig. 1.9 – Représentation des Purines

En raison des propriétés conférées par leur structure chimique, les bases azotées sont des molécules **hydrophobes**. De plus, leur cycle aromatique leur octroie une géométrie essentiellement plane. Cependant, comme pour les sucres, certains atomes les constituant peuvent s'écarter du plan défini par leur noyau aromatique. Or, la diversité de conformations accessibles à la molécule d'ADN n'est pas seulement liée à ces déformations du cycle aromatique mais correspond à la rotation de l'axe défini par la liaison **Base-Sucre** reliant l'atome  $C_{1'}$  des bases azotées à un atome **O** du sucre désoxyribose.

### Le groupement phosphodiester

Une liaison phosphodiester lie un phosphate à deux autres molécules via des liaisons ester. Dans l'ADN, ceci correspond à l'ensemble : **[Sucre-Base]-Phosphate-[Sucre-Base]**, Fig. 1.10. Dans un souci de représentation, nous nous intéresserons au groupement **Phosphate-[Sucre-Base]-Phosphate (P[S-B]P)**. Celui-ci n'est qu'une représentation extrapolée de l'exact groupement phosphodiester. Ce groupement nous permet de discuter l'ensemble des axes de rotation d'un nucléotide, sans oublier l'axe existant entre ce dernier et le phosphate du nucléotide suivant.

Fig. 1.10 – Représentation des axes de rotation d'un groupement  $P[S-B]P$  [Dau93]

Le groupement phosphodiester possède sept angles de rotation indépendants, définis selon la succession d'atomes le constituant :

La chaîne **P-S-P** :

- $\alpha \rightarrow [O-P-O]$
- $\beta \rightarrow [P-O-C]$
- $\gamma \rightarrow [O-C-C]$
- $\delta \rightarrow [C-C-C]$
- $\epsilon \rightarrow [C-C-C-P]$
- $\zeta \rightarrow [C-O-P-O]$

Entre **[S-B]** :

- $\chi \rightarrow [O-C-N-C]$

Les angles de rotation non référencés ci-dessus ne sont pas indépendants. En effet, les valeurs prises par ces derniers sont corrélées aux valeurs des autres angles de la molécule. Ils ne représentent donc pas des variables-clés pour définir la conformation adoptée par le système.

Le **squelette** de l'ADN est constitué d'une succession des groupements **P-S-P**. Ces sept angles de rotation permettent de décrire sa structure. Ces derniers suggèrent de nouveaux degrés de liberté à ajouter à ceux induits par la molécule de sucre ou des bases azotées. Ceci pourrait laisser croire que les conformations accessibles pour la chaîne d'ADN augmentent presque indéfiniment. Or les degrés de liberté associés à des éléments, de façon séparée, ne correspondent pas aux degrés de liberté résultant de ces dits éléments en interaction.

En effet, les interactions du sucre avec les autres composants de l'ADN induisent de nouvelles contraintes. Cet arrangement conduit à de nombreuses restrictions sur la chaîne de l'ADN, limitant ainsi les domaines accessibles à chacun des sept angles de rotation.

Ici, nous avons défini l'angle de torsion selon la succession d'atomes le constituant [A-B-C-D]. Le degré de rotation indiqué correspond à l'angle entre les plans [A-B-C] et [B-C-D], plus facilement visualisable en représentation de Newmann, Fig. 1.11. Les conformations angulaires les plus fréquemment rencontrées sont :

La chaîne **P-S-P** :

- $\rightarrow \sim 60^\circ$
- $\rightarrow \sim 180^\circ$
- $\rightarrow \sim -60^\circ$

Entre **[S-B]** :

- $\rightarrow \sim 0^\circ$
- $\rightarrow \sim 180^\circ$

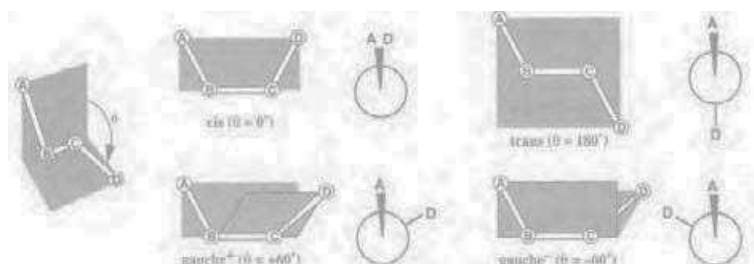


Fig. 1.11 – Représentation des angles de rotation préférentiels [BCT00]

Les contraintes **stériques**, dues à la disposition et au volume occupé par chaque molécule, ainsi que les **barrières de potentiel** de rotation, dues aux énergies requises pour explorer certains états, influencent la conformation de l'ensemble de la chaîne d'ADN. Tout ceci induit des corrélations sur les valeurs de ces angles, limitant ainsi les configurations angulaires accessibles.

### Résumé des corrélations conformationnelles

Nous venons de discuter les conformations préférentielles des constituants d'une chaîne simple brin d'ADN. De plus, nous avons convenu de l'existence de corrélation entre tous ces constituants, même si nous ne les avons pas détaillés.

Le tableau ci-dessous résume les conformations préférentielles résultantes :

Label	Conformation du Sucre	Angle de rotation [S-P] $\Rightarrow \gamma$	Angle de rotation [S-B] $\Rightarrow \chi$
A	C3'-endo	$\sim 60^\circ \rightarrow g^+$	$\sim 180^\circ \rightarrow anti$
B	C2'-endo	$\sim 60^\circ \rightarrow g^+$ $\sim -60^\circ \rightarrow g^-$	$\sim 180^\circ \rightarrow anti$

TABLE 1.1 – Tableau répertoriant les conformations préférentielles, ainsi que leur nom, d'un groupement phosphodiester.

La structure constituée d'un enchaînement linéaire de nucléotide correspond à la **structure primaire** de l'ADN.

### Orientation d'un simple brin

Un brin d'ADN est **orienté**. Cette orientation est définie selon la disposition atomique de la molécule de sucre constituant la chaîne phosphodiester. Examinons de plus près la chaîne [P-S-P] en Fig. 1.10 . Le sucre est lié au phosphate supérieur via son carbone C5' tandis qu'il interagit avec le phosphate suivant via son carbone C3'. L'appellation C5'  $\rightarrow$  C3' désigne ce sens d'interaction défini par orientation de ces deux liaisons. Le sens C5'  $\rightarrow$  C3' définit l'orientation de référence d'un simple brin d'ADN.

*In vivo*, l'existence du caractère orienté de l'ADN est primordial. Cela joue un rôle important dans le bon déroulement de nombreux processus biologiques. En effet, *in vivo* la lecture de l'ADN se fait selon le sens C5'  $\rightarrow$  C3', de même que l'addition successive de nucléotides lors de la synthèse de nouveaux brins d'ADN. La polymérisation de ce brin d'ADN s'effectue par interaction de son extrémité C3' finalisée par un groupement OH libre, avec un nouveau nucléotide constituant un des éléments d'une molécule désoxyribonucléoside triphosphate. La croissance du brin correspond donc à l'allongement de l'ADN dans le sens C5'  $\rightarrow$  C3' par l'ajout successif de groupements à son extrémité C3'.

### 1.2.3 Structures des bases complémentaires

#### Appariement de Chargaff

Les bases azotées s'apparient deux à deux via des **liaisons hydrogènes** selon la complémentarité des paires **A-T** et **G-C**, Fig. 1.12, définies par la **règle d'appariement de Chargaff**. Cette règle de complémentarité permet de respecter les ratios suivants [Cha50] :

- **A=T** : le nombre de bases azotées A est égal au nombre de bases azotées T
- **G=C** : le nombre de bases azotées G est égal au nombre de bases azotées C

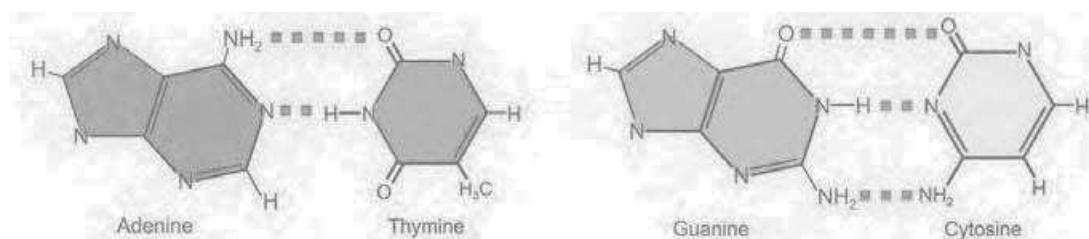


Fig. 1.12 – les appariements de Watson-Crick

Le nombre de bases A (ou T), ou G (ou C) diffère selon les espèces considérées ; par contre, le rapport  $\frac{A+G}{T+C} = 1$  est toujours respecté.

Le nombre de liaisons faibles spécifiques diffère selon la paire considérée. La paire **A-T** est stabilisée par deux liaisons hydrogène tandis que la paire **G-C**, est stabilisée par trois liaisons. L'apport énergétique à fournir pour dissocier, dé-hybrider, de la paire **G-C** est plus important que celle mise en jeu pour la paire **A-T**. En effet, l'énergie de liaison stabilisant la paire **A-T** est de l'ordre de  $-13 \pm 1 \text{ kcal.mol}^{-1}$  soit  $-90 \times 10^{-21} \text{ J}$ ,  $\sim 22$  fois  $E_{\text{them}}$ . L'énergie de  $-26 \pm 2 \text{ kcal.mol}^{-1} \Rightarrow -181 \times 10^{-21} \text{ J}$  permet une stabilisation deux fois plus importante pour la paire **G-C**. Ces valeurs correspondent à la valeur moyenne des énergies d'interactions des couples de bases, obtenues par cinq études *ab initio*, qui sont comparées dans l'article de Sponer *et al.* [SJHo4].

Les premières propositions structurales de l'ADN ne plaçaient pas correctement les groupements hydrogènes mis en jeu dans l'appariement des bases azotées complémentaires [PC53]. Durant un temps, ceci limita la compréhension du ratio  $A=T$  et  $G=C$ . Ce dernier fut pourtant découvert en 1950 par E. Chargaff [Cha50], via la technique de chromatographie sur filtre de papier permettant la séparation des espèces présentes en solution. La pleine compréhension de la spécificité et la complémentarité des bases **A-T** et **G-C** ne fut possible que grâce à la contribution d'autres scientifiques qui élucidèrent la structure exacte des bases azotées. Cette règle de complémentarité permet de respecter le fameux ratio  $A=T$  et  $C=G$  de E. Chargaff. Ceci n'est qu'une anecdote, mais elle rend au puzzle qu'est la recherche un peu de sa complexité et souligne le concours de différentes contributions aux constructions scientifiques.

Il est intéressant d'avoir une idée de la distribution des charges autour des paires de bases. Le schéma suivant, Fig. 1.13, présente l'approximation du potentiel électrostatique de chaque couple **A-T** et **G-C**. Chaque base est mise en interaction avec sa base complémentaire via sa distribution de charge moléculaire où, seules les couches externes sont prises en compte (les couches pleines sont négligées).

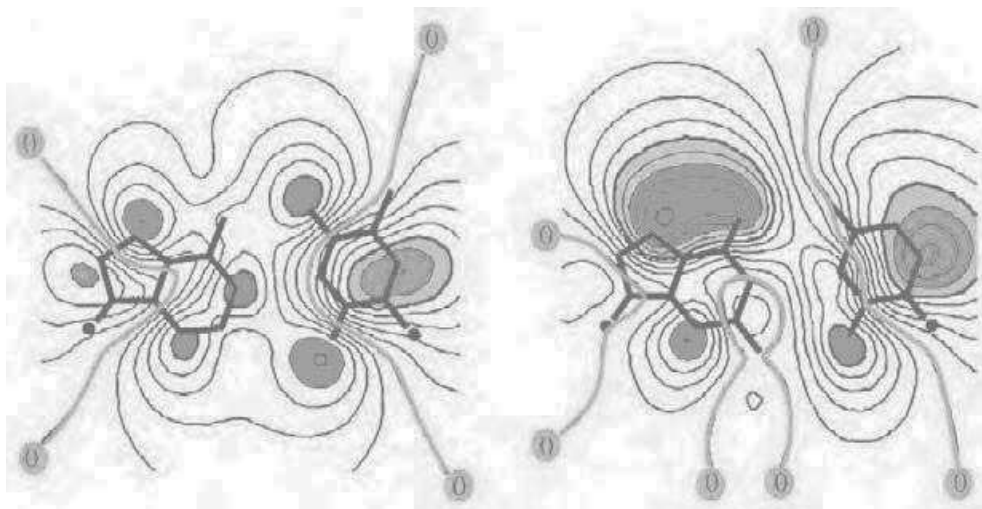


Fig. 1.13 – Représentation de la distribution de charges d'un couple A-T, à gauche et G-C, à droite. Les zones chargées négativement sont représentées en bleu, les zones positives en rouge. Les lignes où le potentiel est nul sont en vert. Adapté de [Hun93]



Le couple **G-C** est associé à une grande zone de charge, par opposition au couple **A-T**. En effet, la base **G** porte une large zone chargée négativement tandis que la base **C** est associée à une large zone positive. Ceci souligne la forte stabilité de ce couple de base au vu de leur interaction. De plus, cette distribution nous informe sur les régions d'interactions préférentielles de chaque couple de base avec des molécules chargées présents dans le milieu.

### Grand et petit sillons

Si nous considérons une paire de bases en interaction avec le squelette [P-S] de l'ADN, les liaisons hydrogène forment un axe de référence. Le point d'attache entre la molécule de sucre et les bases azotées n'est pas symétrique par rapport à cet axe de référence. Dans le plan des paires de base cela conduit à des domaines spatiaux délimités par ces points d'attaches. Ces domaines sont plus grands d'un côté que de l'autre, ce qui correspond respectivement au **grand sillon** et au **petit sillon** de la molécule d'ADN, Fig. 1.14.

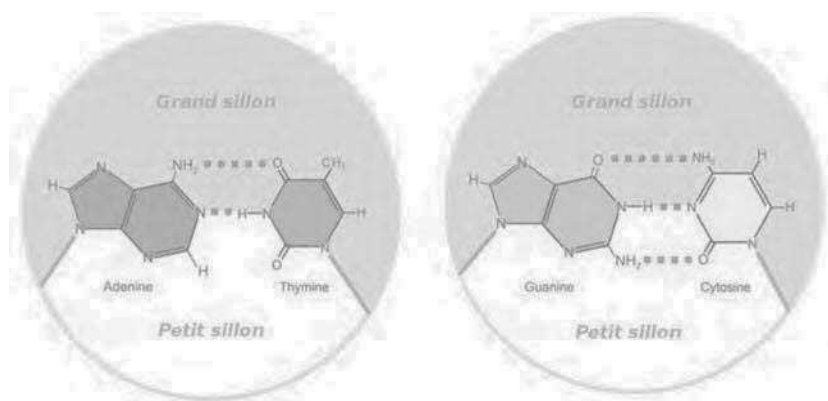


Fig. 1.14 – Représentation du grand et du petit sillon

## 1.3 LA DOUBLE HÉLICE

### 1.3.1 Pourquoi une forme en double hélice ?

Lorsque deux chaînes simple brin sont en interaction, l'agencement de tous les composants de l'ADN conduit à une forme hélicoïdale. Quelles sont les contraintes principales qui induisent cette structure ? Quels sont les principaux arguments de relaxation énergétique de cette structure ?

Pour comprendre simplement cette structure, il existe deux arguments principaux, présentés dans [CD97]. Le **squelette** de l'ADN, constitué de l'enchaînement de molécules de sucre et de phosphates, est essentiellement **hydrophile**. Tandis que les bases azotées, fixées au sucre, sont **hydrophobes**. Les phosphates tendent à adopter un arrangement où ils sont préférentiellement *accessibles* aux molécules d'eau, donc orientés vers l'extérieur de la chaîne. De plus, le volume correspondant à l'encombrement stérique des bases est inférieur à la distance maximale entre les bases. Cette distance maximale correspond à la distance entre l'empilement des paires de bases qui apparaît dans la représentation de la l'ADN, Fig. 1.2 (a). Ici, la formidable flexibilité de conformation de l'ADN entre en jeu.

Les modulations structurales accessibles des chaînes **P-S-P** et **S-B** permettent aux brins d'ADN de s'enrouler l'un autour de l'autre en forme d'hélice sous l'effet des interactions hydrophobes présentes entre les paires de bases. Ainsi, l'espace entre ces paires de base est réduit.

### 1.3.2 Structures d'empilement des bases complémentaires

#### Empilement des bases complémentaires

Une chaîne simple brin d'ADN est constituée d'une succession de nucléotides. La base de chacun de ces nucléotides se lie avec sa base complémentaire afin de former une chaîne double brins. Un double brin d'ADN est composé d'un **empilement** de bases complémentaires qui s'entrelacent selon un agencement en **double hélice**, Fig. 1.2 (b). Les chaînes simple brin la constituant sont antiparallèles, autrement dit les deux brins sont orientés en sens opposés.

La superposition des paires de base conduit à une accumulation du même côté de la double hélice des domaines de grand sillon et du côté opposé des domaines de petit sillon.

L'énergie d'empilement est étudiée via de nombreuses techniques *ab initio*. Durant ces dernières années, grâce au développement des approches de simulation, la précision des valeurs extraites s'est grandement améliorée. Le tableau ci-dessous, [SJM<sup>+</sup>06], répertorie les différentes combinaisons de dimères possédant des énergies d'empilement distinctes. En effet, certaines combinaisons sont de même énergie, elles ne sont donc pas présentes dans le tableau.

Dimère	$E_{\text{empilement}}$ kcal/mol	$E_{\text{empilement}}$ J	Dimère <i>dissymétrique</i>	$E_{\text{empilement}}$ kcal/mol	$E_{\text{empilement}}$ J
$\begin{array}{c} \uparrow A \quad T \downarrow \\ A \quad T \end{array}$	-13.1	$-91 \times 10^{-21}$			
$\begin{array}{c} \uparrow G \quad C \downarrow \\ G \quad C \end{array}$	-11.2	$-78 \times 10^{-21}$			
$\begin{array}{c} \uparrow G \quad C \downarrow \\ T \quad A \end{array}$	-13.4	$-93 \times 10^{-21}$	$\begin{array}{c} \uparrow T \quad A \downarrow \\ G \quad C \end{array}$	-15.1	$-105 \times 10^{-21}$
$\begin{array}{c} \uparrow G \quad C \downarrow \\ A \quad T \end{array}$	-12.9	$-90 \times 10^{-21}$	$\begin{array}{c} \uparrow A \quad T \downarrow \\ G \quad C \end{array}$	-12.5	$-87 \times 10^{-21}$
$\begin{array}{c} \uparrow T \quad A \downarrow \\ A \quad T \end{array}$	-12.8	$-89 \times 10^{-21}$	$\begin{array}{c} \uparrow A \quad T \downarrow \\ T \quad A \end{array}$	-13.3	$-92 \times 10^{-21}$
$\begin{array}{c} \uparrow G \quad C \downarrow \\ C \quad G \end{array}$	-15.8	$-110 \times 10^{-21}$	$\begin{array}{c} \uparrow C \quad G \downarrow \\ G \quad C \end{array}$	-17.3	$-120 \times 10^{-21}$

TABLE 1.2 – Tableau répertoriant les valeurs des énergies d'empilement d'après [SJM<sup>+</sup>06].

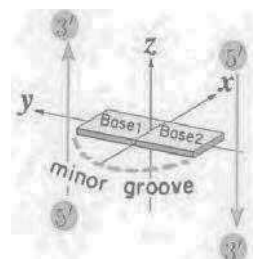
Sans entrer plus amplement dans les détails du calcul, nous pouvons constater que l'énergie d'empilement est dépendante de la séquence même de l'ADN. En effet, l'énergie d'un dimère  $\begin{array}{c} \uparrow G \quad C \downarrow \\ T \quad A \end{array}$  est de l'ordre de  $\sim 22 \times E_{\text{them}}$  tandis que son dimère *dissymétrique*  $\begin{array}{c} \uparrow T \quad A \downarrow \\ G \quad C \end{array}$  possède une énergie de  $\sim 25 \times E_{\text{them}}$ . Ceci attire notre attention sur l'importance de la séquence de l'ADN.

Cette structure hélicoïdale permet un empilement des paires de bases et minimise l'énergie structurale mise en jeu. L'énergie d'interaction de deux bases complémentaires dépend des constituants du système étudié. L'énergie d'empilement quant à elle, dépend aussi des composants mais surtout de la composition des couples de bases proches voisins, autrement dit de la séquence de l'ADN. Afin de minimiser les contraintes, la forme stable de l'ADN double brin est une double hélice.

Cette représentation sous forme de double hélice correspond à la **structure secondaire** de l'ADN.

### Modulation angulaire autour des bases complémentaires

Le colloque de Cambridge en 1988, a permis de standardiser la notation ainsi que la nomenclature d'un grand nombre de paramètres hélicoïdaux [Dic89]. Pour décrire les modulations et effets de structure observables dans l'ADN, nous utilisons le référentiel local défini ci-contre.



Le panel de paramètres hélicoïdaux représentés en Fig. 1.15, définit les modulations structurales de **translation** et **rotation** accessibles dans un double brin d'ADN. Ils caractérisent le positionnement intra-paire de base (des bases entre elles) ou inter-paire de base (variation de la topologie de l'empilement des paires des bases) ou bien encore leur positionnement par rapport à l'axe de l'hélice. Ces paramètres rendent compte de la flexibilité de la structure de la biomolécule.

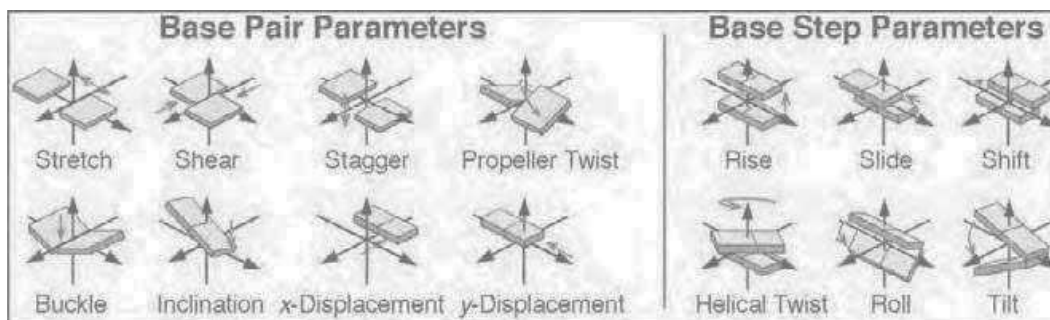


Fig. 1.15 – Représentation des différentes modulations structurales possibles pour les bases appariées ainsi que l'empilement des paires de bases appariées. Schéma adapté de [BCToo].

### Orientation de la double hélice

Le sens d'enroulement des deux brins d'ADN est important pour décrire les structures en double hélice. Une façon simple de distinguer ces deux sens de rotation, est de se représenter tenant l'hélice dans notre main. L'hélice sera **droite** si, en montant selon la direction du pouce, elle tourne selon l'orientation des doigts de votre main droite. Ou bien, elle sera **gauche** si elle tourne dans la direction des doigts de votre main gauche.

Les termes **horaire** et **anti-horaire** sont aussi utilisés pour définir le sens de rotation de l'hélice. Une hélice est droite si, lorsque nous avançons selon son axe tout en l'éloignant de nous (observateur), nous pouvons voir la trajectoire de ces

brins tourner selon le sens horaire (sens des aiguilles d'une montre). Une hélice gauche tourne selon le sens **anti-horaire**.

### 1.3.3 Les formes principales de double hélice

Il existe trois grandes structures de double hélice dissociables de par leurs paramètres structuraux et dont les représentations sont regroupées en Fig. 1.16. Celles-ci découlent des premières interprétations des diagrammes de diffraction de rayons X par des fibres d'ADN.

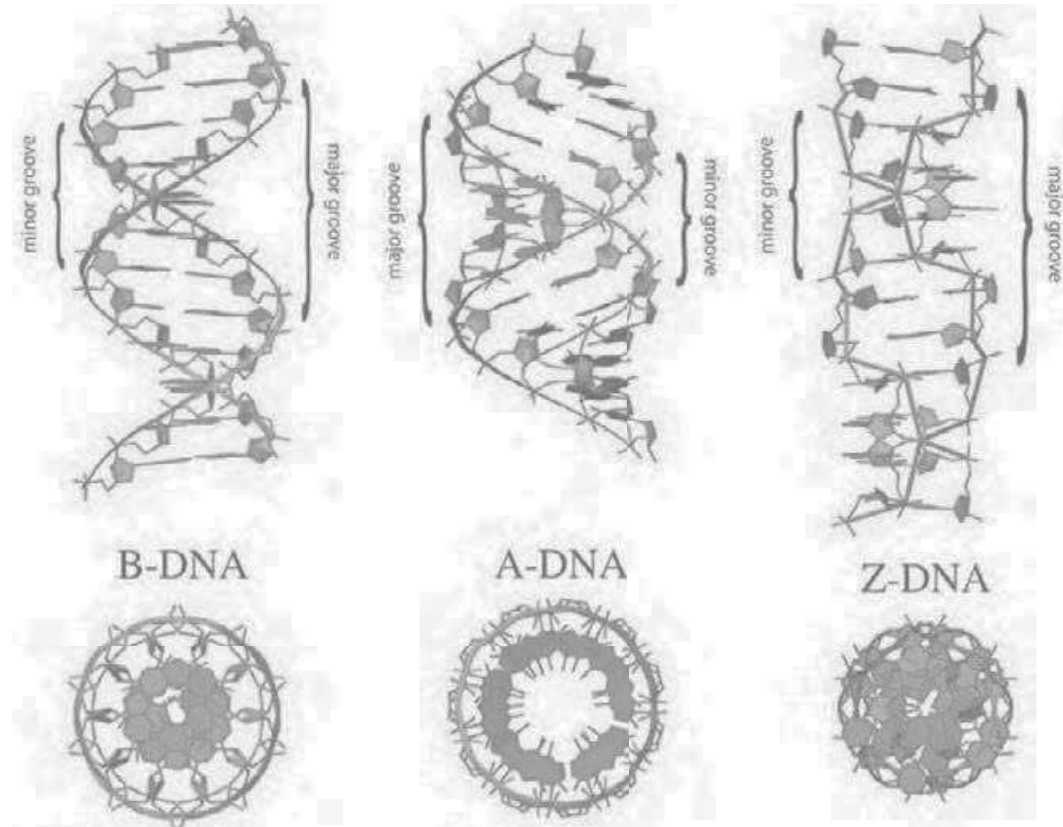


Fig. 1.16 – Représentation des différentes structures de l'ADN. Pour chacune des formes, une vue de profil et une vue selon l'axe des hélices, sont présentées [TV]. Les phosphates sont représentés par un trait. Les termes anglais major groove et minor groove désignent respectivement le grand et petit sillon.

Les deux premières formes : **A** et **B**, furent initialement découvertes grâce aux travaux de M. H. F. Wilkins *et al.* [WSW53] ainsi que Franklin et R. G. Gosling, [FG53a, FG53b]. La **forme B** correspond à la fameuse structure proposée par les scientifiques J. D. Watson et F. Crick [WC53] afin d'interpréter ces résultats expérimentaux [FG53a]. La **forme A**, dont les clichés furent aussi réalisés en 1953 par R. E. Franklin et R. G. Gosling [FG53b], se différencie par une structure, proposée par H. R. Drew *et al.* [DWT<sup>+</sup>81], plus compacte selon l'axe de la double hélice.

La **forme Z**, quant à elle, possède une structure nettement dissociable des précédentes, ceci s'explique par les contraintes présentes dans sa double hélice. Cette dernière fut, en premier lieu, proposée dans les papiers suivants [WQK<sup>+</sup>79, DTT<sup>+</sup>80].

Les nombreuses études cristallographiques qui suivirent, ont permis de valider, raffiner et préciser les détails topologiques de ces différentes structures de double hélice [MSWH58, DWT<sup>+</sup>81].

Les paramètres associés aux différentes structures tridimensionnelles de l'ADN que nous allons décrire, ne sont en réalité, que des valeurs moyennes. En effet, comme nous avons pu le constater depuis le début de ce manuscrit, il existe des modulations structurales fines qui sont dépendantes des éléments constituant la séquence de l'ADN. Malgré les contraintes et restrictions sur la structure de l'ADN, celui-ci conserve une relative flexibilité.

### Forme canonique B

Cette structure, Fig. 1.17, est une double **hélice droite** formée de deux chaînes phosphodiester **antiparallèles** : l'une possède une orientation  $5' \rightarrow 3'$  dans le sens opposé de sa chaîne complémentaire. Le sucre désoxyribose adopte la géométrie  $C_{2'}$  **endo** [Dic92], l'une des conformations préférentielles définie en Table 1.1.

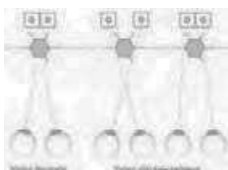
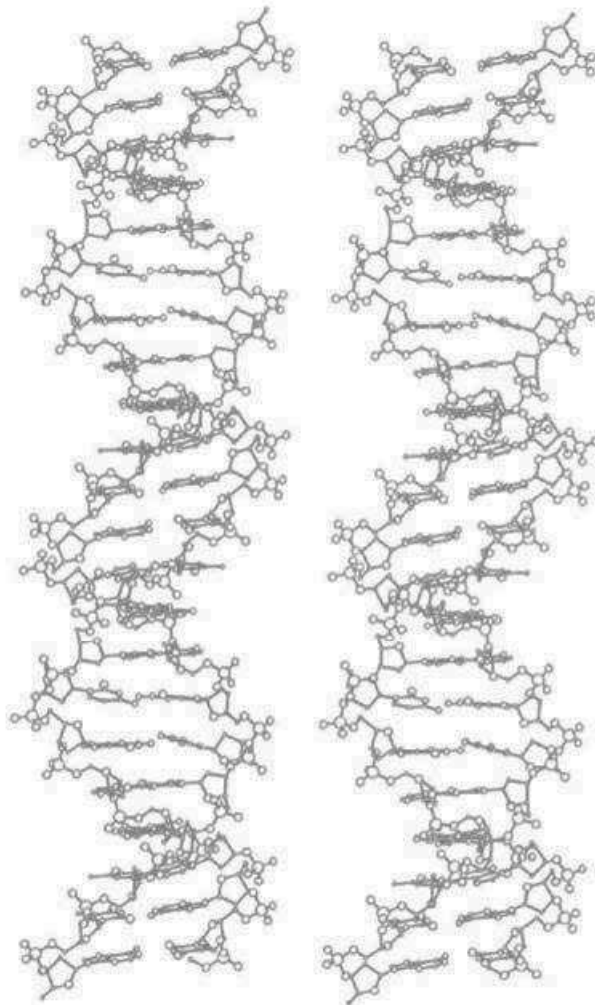


Fig. 1.17 – Structure B de la double hélice d'ADN. Représentation extraite de [Dic92]. Cette représentation peut-être visualisée en 3 dimensions en vision stéréogramme, pour ceux qui y réussissent. La possibilité de vision 3D sera signalée par le logo ci-contre

L'enchaînement des molécules phosphate-sucre est régulier. Les groupements phosphates sont orientés vers l'extérieur de la double hélice. Ainsi, ils contribuent à la solvation de la molécule d'ADN en milieu aqueux, [Her93]. En effet, les phosphates s'ionisent, conférant un caractère **chargé négativement** à la molécule. Le groupement phosphodiester possède un  $pK_a \sim 1$ , [NCK<sup>+</sup>93], plus la valeur est faible plus la force de l'acide est forte. De plus, ces derniers, étant excentrés vers l'extérieur de l'objet, sont plus disponibles. Les groupements phosphates sont accessibles à des interactions avec des molécules d'eau et/ou des ions en solution. Les bases sont perpendiculaires à l'axe de l'ADN. De plus, une base **A** n'est pas distinguable d'une base **C**, de par sa géométrie d'accroche au squelette de l'ADN. L'axe de l'hélice se situe au centre de l'appariement de ces bases complémentaires.

L'empilement des bases complémentaires détermine un tour d'hélice d'une longueur d'environ 3.4 nm, soit une moyenne de 10.5 paires de bases par tour d'hélice. Ceci mène à la fameuse loi de conversion  $1 \text{ bp} \iff 0.34 \text{ nm}$ , [Ast39].

Autrement dit, un tour d'hélice de 10.5 bp correspond à environ 21 nucléotides pour un diamètre de l'ordre de 2.0 nm. Le petit et le grand sillon de l'ADN sont nettement visibles et distinguables sur cette **forme B**. La largeur typique du grand sillon est de l'ordre de 1 nm, presque deux fois plus importante que le petit, [WDT<sup>+</sup>80]. Cette forme d'ADN, par rapport aux autres formes principales, possède la plus importante taille de grand sillon. Cela favorise l'accessibilité des bases azotées pour la formation de complexe ADN-protéine ou l'intercalage de molécule, [AE97]

La **forme B** apparaît être la structure la plus stable en solution aqueuse. Elle correspond à la conformation typique de l'ADN *in vivo*, lorsque l'ADN est soit associé à des protéines, soit sous forme de chromatine, [RZ03].

### Forme canonique A

Là encore, nous sommes en présence d'une **hélice droite** formée de chaînes simple brin antiparallèles, Fig. 1.18. Cette fois-ci, le sucre adopte une configuration **C<sub>3'</sub> endo** [Dic92], son cycle diffère alors par rapport à la forme **B**.

Le squelette phosphaté est moins régulier que dans la phase B. Les groupements phosphates sont repoussés vers le grand sillon. Ainsi, ils sont moins disponibles pour interagir avec les molécules d'eau du solvant.

Cette structure diffère essentiellement de la forme **B** du fait de la position des paires de bases, facilement identifiables sur la vue selon l'axe de l'hélice en Fig. 1.16. L'axe de l'hélice ne passe plus par le centre des paires de bases. Les bases sont localement perpendiculaires à l'axe de l'ADN, mais forment un angle de  $19^\circ$  avec l'axe global de l'hélice. Les bases sont repoussées vers la périphérie de l'ADN.

L'empilement des bases complémentaires confère à l'ADN **A** un caractère plus compact suivant l'axe de la molécule. En effet, dû à l'orientation des phosphates vers le grand sillon, ce dernier en devient plus étroit et plus profond. En revanche, le petit sillon est plus large et moins profond. Un tour d'hélice correspond à une longueur d'environ 2.9 nm, soit une moyenne de 11 paires de bases.

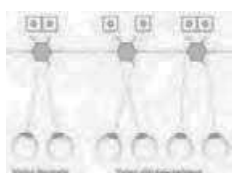
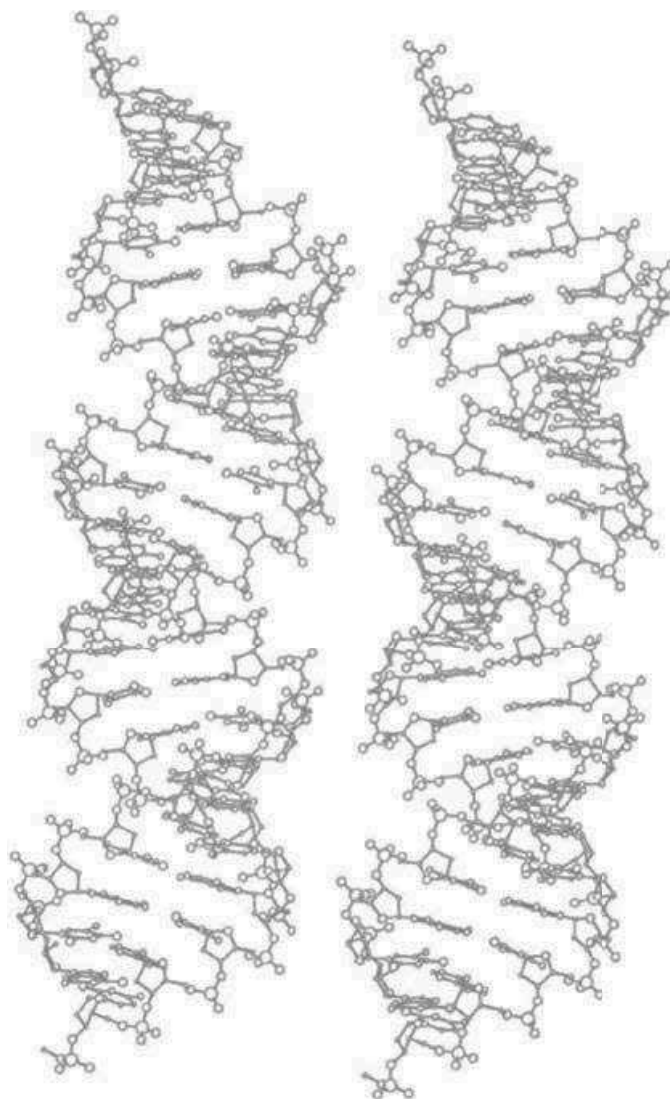


Fig. 1.18 – Structure A de la double hélice d'ADN. Représentation extraite de [Dic92].

La forme A est peu répandue *in vivo*, où elle est essentiellement observée sur les systèmes ARN/ADN lors de la transcription. Elle s'observe sur des échantillons d'ADN présentant des conditions d'hydratation plus faibles [Her93], par exemple en présence d'éthanol. Cette forme fut obtenue pour la première fois sur des fibres d'ADN déshydratées observées par rayon X, lorsque l'humidité relative est diminuée.

### Forme Z

Cette forme Z est composée d'une **hélice gauche**, Fig. 1.19. L'unité structurale est formée de deux paires de bases, et non plus d'une, comme pour la forme **B** et **A**. Le sucre alterne entre les conformations :  $C_3'$  **endo** relié aux bases **G** et  $C_2'$  **endo** relié aux bases **C**.

De plus, les groupements phosphates se répartissent alternativement dans le petit et le grand sillon. Cette forme en zigzag du squelette phosphaté lui a valu son nom. Cette répartition des phosphates déstabilise cette structure, il en résulte des répulsions électrostatiques. S'ajoutant à la contrainte sur les sucres, celle-ci entraîne la formation locale d'un angle de  $9^\circ$  entre les bases et à l'axe de la molécule d'ADN. Les paires de bases ont tendance à être repoussées vers l'extérieur de l'ADN. En d'autres termes, l'axe de l'hélice s'écarte significativement des paires de bases.

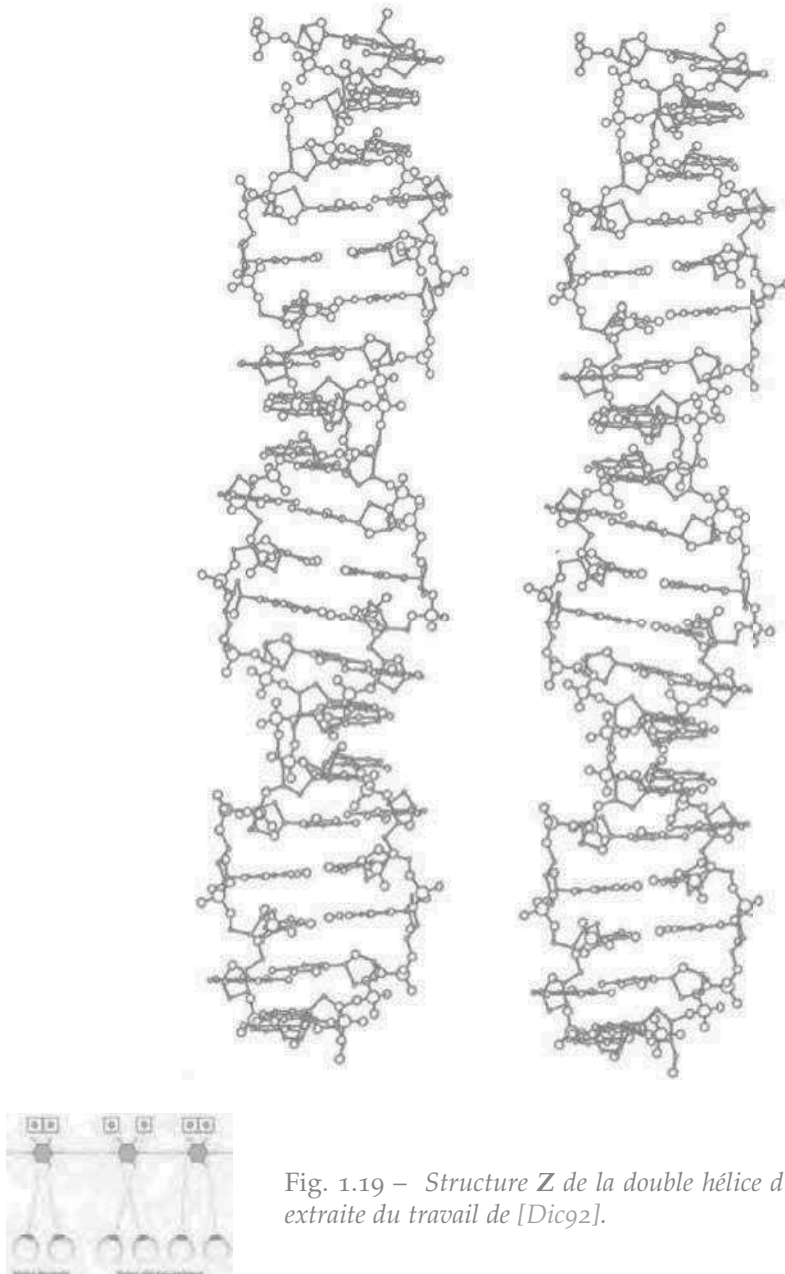


Fig. 1.19 – Structure Z de la double hélice d'ADN. Représentation extraite du travail de [Dic92].

Cette double hélice est plus étroite. De plus, elle est plus allongée afin de minimiser la répulsion des charges des phosphates voisins. Cette hélice possède un diamètre d'environ 1.8 nm pour un tour d'hélice d'environ 4.5 nm. Ceci correspond à 12 paires de bases par tour d'hélice. De par la répartition des phosphates, le grand sillon a disparu. Cette structure n'a donc qu'un sillon profond, qui peut être référencé comme étant le petit sillon.



La **forme Z** est une structure de transition, elle s'observe sur des courtes séquences d'ADN et préférentiellement dans des **régions riches en paires G-C**. Vargason *et al.* ont montré la capacité de transition de l'ADN d'une structure de **forme B** à une **forme A** grâce à la présence d'une séquence (GGCGCC) [VHH01]. En particulier, Strick *et al.* [SAB<sup>+</sup>96] ont montré, sur une même molécule d'ADN soumise à une contrainte de torsion et d'étirement, la coexistence de différentes formes d'hélice. Ainsi l'ADN adopte la **forme Z** pour relaxer cette contrainte, plutôt que l'adoption d'une autre forme intermédiaire.

De plus, de nombreuses études ont montré des interactions spécifiques entre protéines et des séquences de **forme Z**, ce qui sous-entend un rôle spécifique lors de processus biologiques [RZ03]. Le phénomène de méthylation correspond à l'ajout d'un groupement méthyl-CH<sub>3</sub> sur le carbone C'<sub>5</sub> des cytosines lorsqu'elles sont associées à un doublet  $\begin{array}{cc} \uparrow C \text{ ou } G & \dots \\ G \text{ ou } C & \dots \end{array}$ . Ce phénomène joue un rôle dans la régulation des gènes [HR96]. Bien souvent, ces zones sont synonymes de peu d'expression des gènes, ceci est lié à la difficulté de fixation des protéines responsables de la transcription. Dans ces régions l'ADN adopte une **forme Z** [TDB<sup>+</sup>12].

### Principales variations des paramètres structuraux

Nous venons de constater que les valeurs de certains paramètres structuraux permettent de définir les différentes topologies structurales accessibles à la molécule d'ADN. Nous pouvons ainsi résumer les caractéristiques des principales formes de l'ADN dans le tableau ci-dessous.

Caractéristique	Forme B	Forme A	Forme Z
Spécificité de la séquence	non spécifique	non spécifique	riche en C,G
Orientation de l'hélice	droite	droite	gauche
Unité structurale	1 bp	1 bp	2 bp
Conformation du sucre	C <sub>2'</sub> -endo	C <sub>3'</sub> -endo	C → C <sub>2'</sub> -endo G → C <sub>3'</sub> -endo
Distance axiale entre les bases	0.23 nm	0.23 nm	0.38 nm
Diamètre de l'hélice	2.4 nm	2.6 nm	1.8 nm
Angle de basculement des bases par rapport à l'axe de l'hélice	~ 0°	~ 19°	~ -9°
Nombre de pb par tour d'hélice	10.5 pb	11 pb	12 pb
Pas de l'hélice	3.4 nm	2.9 nm	4.5 nm
Paramètres du Grand sillon :			
Largeur	1.17	0.27	∅
Profondeur	0.85	1.35	∅
Paramètres du Petit sillon :			
Largeur	0.57	1.10	0.40
Profondeur	0.75	0.28	0.90

TABLE 1.3 – Comparaison des propriétés structurales des ADNs A,B et Z, d'après [Dic92, Dic99].

De plus, il existe d'autres formes structurales accessibles à la molécule d'ADN. Celles-ci semblent correspondre à des structures intermédiaires de l'ADN lors d'événements bien spécifiques intervenants *in vivo*.

L'ADN peut adopter une forme en *épingle à cheveux* due à des interactions intra-brin formant un repliement. Ce type de repliement peut aussi exister sur les deux simples brins en parallèle, la structure résultante est ainsi cruciforme. Ces formes sont localisées sur certaines séquences présentant des répétitions inversées qui permettent ce type de repliement.

La forme H ou triplex est aussi localisée sur certaines séquences de l'ADN. Cette structure correspond à une ouverture locale de la double hélice, sur laquelle un simple brin vient s'insérer et s'apparier. Dans cette région, la double hélice est en interaction avec un troisième brin d'ADN formant une triple hélice.

L'ADN ramifié, quant à lui, peut être observé lorsque les deux simples brins cessent d'être en interaction ; vient alors s'insérer un troisième simple brin d'ADN. Ce type de ramification peut s'ajouter en série, donnant ainsi naissance à une structure aux ramifications multiples.

#### 1.4 LA NOTION DE FLEXIBILITÉ DE L'ADN DANS LA CELLULE

*In vivo*, une molécule d'ADN n'est jamais *nue*. Un grand nombre de molécules telles que des protéines, des enzymes sont *fixées* le long de ce polymère. Elles ont pour effet, entre autres, de modifier localement les propriétés physiques du polymère.

##### 1.4.1 La compaction de l'ADN

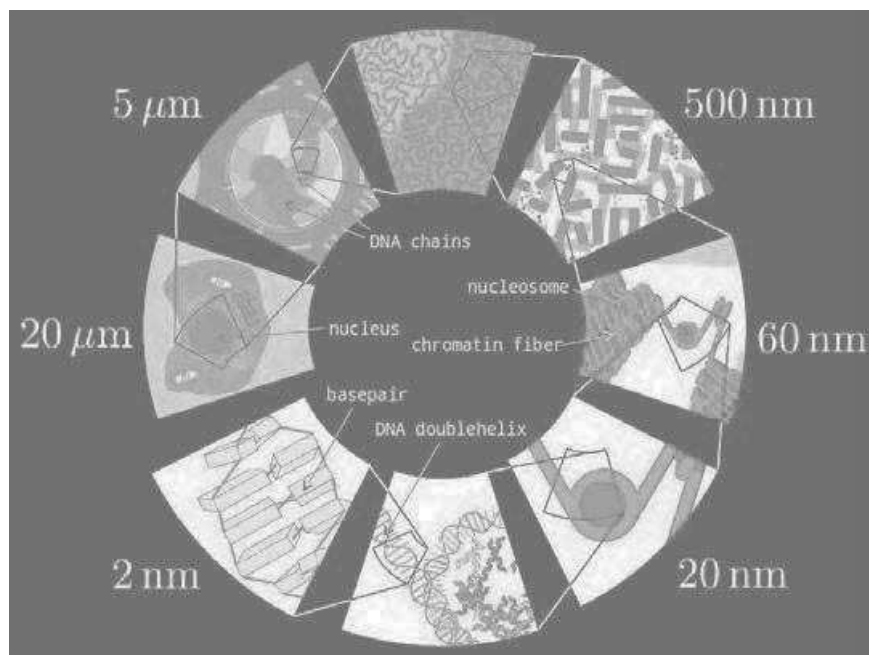


Fig. 1.20 – Vision schématique de la molécule d'ADN à différentes échelles d'observation. Schéma extrait du cours d'Helmut Schiessel, Les Houches 2013 ou [Sch14]

Dès le collège, les professeurs enseignent que l'information génétique est contenue dans le noyau cellulaire pour les organismes eucaryotes<sup>4</sup>, notamment chez l'Homme. Puis, une des autres informations principales communiquées est que "l'ADN est sous forme de chromosomes". Or, la forme chromosomique n'est en fait, qu'un pâle aperçu du large panel des formes possibles. En effet, la molécule d'ADN possède une riche palette de structures accessibles.

Le diagramme en Fig. 1.20, nous présente une vue de l'ADN chez un eucaryote. Ainsi, nous passons de l'échelle microscopique dans le noyau cellulaire, dessin à l'échelle de 20  $\mu\text{m}$  milieu gauche, puis par l'enroulement de l'ADN autour des nucléosomes, dessin à l'échelle de 60 nm milieu droit, jusqu'au détail de la structure en double hélice de l'ADN, distinguable à l'échelle nanoscopique.

Nous pouvons donc nous intéresser à l'ADN à différentes échelles d'observation. Il peut être extrêmement bien ordonné et compacté sous sa forme la plus célèbre qu'est le **chromosome** ou encore être partiellement *rangé* sous forme de **pelote aléatoire**. De plus, le détail de ces structures s'éclaire en fonction de l'échelle à laquelle nous les observons.

Ceci nous permet de mettre en évidence différentes conformations adoptées par notre génome, et ainsi, de mesurer la grande variabilité et diversité de la structure que peut adopter la molécule d'ADN.

### 1.4.2 La notion de flexibilité

Notre génome est constitué de 26 paires de molécules d'ADN ou chromosomes. Sa correspondance en nombre de paires de bases est de  $N_{\text{Génome}} = 3.10^9$  bp et ce, en deux exemplaires dans les cellules somatiques<sup>5</sup>. Autrement dit 6 Gbp sont confinées dans le noyau cellulaire.

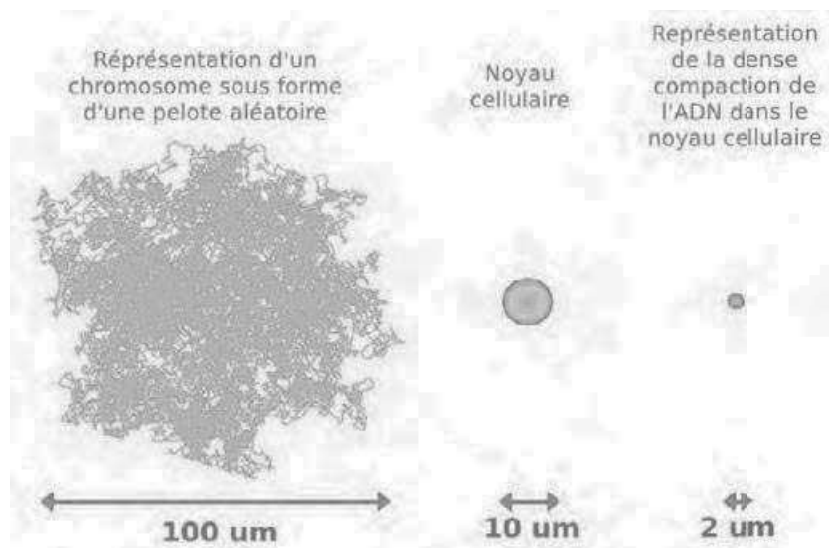


Fig. 1.21 – Représentation de la capacité de compaction de l'ADN dans la cellule

4. Eucaryotes : organismes qui se caractérisent par la présence d'un noyau. En opposition aux organismes procaryotes qui se caractérisent par l'absence de noyau.

5. Cellules somatiques : cellules qui ne sont pas à l'origine de gamètes, par opposition aux cellules germinales.

La longueur de contour  $L$ , correspondant à l'échelle de taille la plus simple d'une molécule d'ADN, est  $L = N_{\text{Gnome}} \times 0.34 \text{ nm}$ . Cela implique une *représentation* linéaire pour notre génome égale à une molécule d'ADN étirée d'environ 2 m. L'aspect le plus stupéfiant est que ces 2 m d'ADN sont condensés au sein du noyau cellulaire qui possède un rayon typique d'environ  $5 \mu\text{m}$  !

Si nous essayons de la ranger sous la forme de **pelote aléatoire**, une molécule d'ADN correspond à une sphère d'environ  $100 \mu\text{m}$  de rayon, Fig. 1.21. Là encore, la molécule d'ADN est spatialement plus importante que le noyau cellulaire. Il est donc nécessaire d'introduire une notion de flexibilité pour permettre la compaction de cette molécule. Il est alors possible de stocker cette molécule compactée, ainsi que les 45 autres chromosomes présents, dans le noyau.

La prise en compte de la flexibilité de la molécule d'ADN permet et explique, en partie, son assemblage en chromatine qui conduit à sa compaction dans de tels petits volumes. Cette flexibilité est liée à la grandeur caractéristique qu'est la longueur de persistance. Nous la définirons dans les parties 2.2.3 et 2.3.1.

## CONCLUSION DU CHAPITRE

La large gamme de conformations et d'angles de rotations accessibles pour les constituants de l'ADN, lui apporte une grande flexibilité. Par ailleurs, les corrélations conformationnelles entre ces constituants ainsi que les barrières énergétiques induisent une certaine rigidité à cette biomolécule. L'ADN est donc un objet semi-flexible possédant une structure non statique, comme l'existence de nombreuses topologies de la double hélice a pu le souligner.

Nous venons de voir que les formes principales de l'ADN ne se rencontrent pas exactement dans les mêmes conditions. L'existence de ces trois hélices souligne l'importance de deux grands paramètres influant sur la conformation structurale de l'ADN : **l'environnement local** et **la séquence même de l'ADN**.

En effet, la **forme A** s'observe dans des conditions d'hydratation plus faibles [Her93]. Le changement de géométrie est influencé par les interactions et la capacité d'interactions avec les molécules d'eau. **L'environnement local** et la topologie adoptée par l'ADN sont fortement corrélés. La structure de l'ADN **Z**, quant à elle, est le résultat de contraintes structurales fortes, induites par la répétition des bases **G** et **C**. Afin d'acquiescer la stabilité de l'objet, les sucres adoptent une conformation correspondant à une alternance successive de leur topologie. Ceci souligne l'importance de **la séquence même de l'ADN**, et offre un magnifique exemple, de la capacité de modulation de la structure de l'ADN.

Tout ceci finit de nous convaincre du caractère flexible et dynamique de la molécule d'ADN. Cette biomolécule répond aux exigences locales par une modulation et/ou déformation de ses paramètres structuraux à plus ou moins grande échelle, afin de relaxer les contraintes.

Depuis de nombreuses décennies l'importance de la séquence a été prouvée. Celle-ci n'est pas seulement cruciale pour le contenu de l'information elle-même, elle est aussi essentielle et déterminante pour, par exemple, la fixation d'une protéine sur un site. Les différences induites par la séquence ont, entre autres, une traduction et une implication sur la distance entre les bases successives. L'énergie mise en jeu pour deux **A** successives ne correspond pas à celle pour deux **G** successives. Il en est de même pour l'énergie d'interaction entre un couple **A-T** ou **C-G**. L'énergie thermique nécessaire à la dissociation de bases complémentaires implique une température de dissociation plus élevée pour **G-C** que pour **A-T**. Ceci implique une différence comportementale importante qui est profitable lors des processus biologiques. Les étapes de réplication et traduction exploitent ces différences afin de créer l'ouverture d'une bulle de dénaturation dans la double hélice d'ADN.

Quels sont donc nos outils pour décrire l'ADN ? Cet objet est souvent défini par son comportement *global*, qui correspond à l'échelle d'observation mésoscopique. Cette échelle intermédiaire permet son étude entre l'échelle microscopique, caractérisant les atomes le constituant, et l'échelle macroscopique, caractérisant la molécule d'ADN dans son ensemble. Or nous, l'avons souligné, il ne peut se résumer à une simple chaîne homogène. Jusqu'où pouvons nous décrire son comportement local ? Dans le chapitre suivant, nous allons entamer la description des modèles physiques qui décrivent les propriétés de la molécule d'ADN.



# LE POLYMÈRE ADN

# 2

## SOMMAIRE

INTRODUCTION . . . . .	34
2.1 MODÈLES STATISTIQUES DE POLYMÈRE . . . . .	35
2.2 MODÈLES FLEXIBLES . . . . .	35
2.2.1 Modèle de la chaîne gaussienne sur un réseau . . . . .	35
Probabilité . . . . .	36
2.2.2 Modèle de la chaîne librement jointe : <i>Freely Jointed Chain (FJC)</i> . . . . .	38
Longueur de contour : $L$ ( <i>en : contour lenght</i> ) . . . . .	38
Distance bout-à-bout : $R$ ( <i>en : end-to-end distance</i> ) . . . . .	38
Rayon de giration : $R_g$ ( <i>en : radius of gyration</i> ) . . . . .	39
L'entropie : $S$ . . . . .	40
2.2.3 Limite des modèles flexibles pour décrire une chaîne réelle . . . . .	41
Longueur statistique de Kuhn : $b$ ( <i>en : Kuhn lenght</i> ) . . . . .	41
2.3 MODÈLES SEMI-FLEXIBLES . . . . .	42
2.3.1 Modèle de Kratky-Porod (KP) ou <i>Freely Rotating Chain (FRC)</i> . . . . .	42
Distance bout-à-bout . . . . .	42
Longueur de persistance : $L_p$ ( <i>en : persistence lenght</i> ) . . . . .	43
2.3.2 Modèle du "Ver" : <i>Worm Like Chain (WLC)</i> . . . . .	43
Longueur de persistance . . . . .	43
Module de courbure . . . . .	44
Distance bout-à-bout et Longueur de persistance . . . . .	44
Contribution des modèles semi-flexibles . . . . .	45
2.3.3 Limite des modèles semi-flexibles pour décrire une chaîne réelle . . . . .	45
2.4 CHAÎNE RÉELLE . . . . .	46
2.4.1 Théorie de Flory . . . . .	46
Effet de volume exclu . . . . .	46
Gonflement de la chaîne . . . . .	47
Cas de la molécule d'ADN . . . . .	47
Bon et mauvais solvant . . . . .	48
CONCLUSION . . . . .	49

## INTRODUCTION

Il existe un grand nombre de **polymères** dans la nature. Un **polymère** est une molécule constituée d'une succession d'unités élémentaires semblables, nommées **monomères**, connectées les unes aux autres. Le bio-polymère le plus connu reste l'ADN, molécule d'intérêt de cette thèse. Comme nous venons de le voir dans le chapitre précédent, il peut être sous la forme de **simple brin (ss)**, (*en : single stand*), ou de **double brin (ds)**, (*en : double stand*), correspondant à un appariement des **bases azotées complémentaires**, ou non. De plus, ce système, de par sa **flexibilité structurale**, possède un grand nombre de conformations accessibles. Ces conformations restent toutefois **limitées à des domaines de variation** dus aux angles constituant la chaîne de nucléotides, ce qui confère à l'ADN une certaine **rigidité**. La description physique des **conformations géométriques** accessible à un tel bio-polymère est un exercice intéressant et important. Ainsi, un grand nombre de **modèles physiques** ont été développés pour permettre la description, la compréhension et la caractérisation de la **structure tridimensionnelle** de ce type de système, qu'il soit sous forme de **chaîne** simple ou double brins.

*In vivo*, la molécule d'ADN est en solution et n'est jamais isolée. Autrement dit, elle est en contact perpétuel avec des molécules d'eau du solvant, des ions ou d'autres bio-molécules. Ces contacts peuvent être de longue durée si ce sont des interactions physico-chimiques, ou des contacts instantanés dus à de simples collisions. Toutes ces **collisions aléatoires** contribuent à l'exploration du **domaine conformationnel du polymère**. La conformation de la molécule d'ADN est ainsi, à l'échelle temporelle d'observation de l'Homme (de l'ordre de  $\mu$ s, ms ou même s), en perpétuelles fluctuations dues aux **mouvements browniens**. L'essentiel des propriétés de notre objet peut être appréhendé à l'aide de quelques modèles simplifiés. Ces modèles offrent une **description analytique** via les méthodes de la **physique statistique**. Ainsi, les phénomènes et propriétés de diffusion de notre objet sont transcrites en termes de **probabilités statistiques**.

Ce chapitre recense donc, les principaux **modèles statistiques** utilisés afin de décrire un polymère libre en solution. Ainsi étudiée, la conformation de l'ADN dépend de la **probabilité statistique** d'existence de sa géométrie conformationnelle. Nous allons voir qu'il existe des modèles physiques à plusieurs degrés de complexité, pouvant décrire l'ADN. Le modèle le plus simple ne prend pas en compte les corrélations existantes entre les différentes sous-parties de la molécule. Les modèles de polymères semi-flexibles prennent en compte ces corrélations sous un terme d'élasticité correspondant à une flexibilité réduite qui conduit à une certaine rigidité de la molécule. Enfin, nous verrons qu'il est aussi possible d'ajouter des paramètres effectifs pour tenir compte, au moins partiellement, de la contribution du solvant. Nous verrons ensuite que pour décrire les **propriétés dynamiques** de tels systèmes il est nécessaire de raffiner ces modèles statistiques en y ajoutant des interactions hydrodynamiques.



## 2.1 MODÈLES STATISTIQUES DE POLYMÈRE

Un **polymère**, à strictement parler, est une molécule constituée d'une succession de  $N$  **monomères**, l'unité semblable, connectés entre eux par des **liaisons covalentes**. Le nombre  $N$  de monomères correspond au **degré de polymérisation** de la chaîne. L'unité semblable de l'ADN à l'échelle minimale est le nucléotide pour une chaîne simple brin, ou une paire de nucléotides complémentaires pour une chaîne double brins. Nous parlons de polymère, ou d'hétéro-polymère, même si les nucléotides ne sont pas exactement identiques, en terme de séquence. Pour travailler à une échelle supérieure, nous pouvons considérer un tour d'hélice, soit 10.5 nucléotides pour un brin d'ADN de la forme **B**, comme notre unité élémentaire. Cela sous-entend qu'une partie de la structure fine de l'ADN n'est pas utilisée dans cette échelle de description des **propriétés statistiques** du polymère. Dans ce chapitre nous allons nous intéresser à la description théorique des polymères et aux paramètres clés, mécaniques et structuraux, qui le décrivent.

L'analyse conformationnelle d'une chaîne de polymère **isolée** dépend de la prise en compte d'interactions à deux niveaux distincts.

- Tout d'abord, il existe des **interactions à courte distance** qui définissent le caractère flexible et/ou rigide de la chaîne, comme nous venons de les répertorier précédemment. Ces interactions sont relatives aux angles de rotation ainsi qu'aux contraintes stériques induites entre les différents éléments consécutifs qui constituent le squelette de l'ADN. Les interactions électrostatiques y jouent aussi un rôle essentiel.
- Il existe aussi des **interactions à longue distance**. Ces dernières sont directement liées aux **effets de volume exclu** qui interdisent à tout segment de la chaîne de revenir en un point déjà occupé par ladite chaîne. Comme nous le verrons plus tard, l'**effet de volume exclu**, répulsion entre sous-parties de la chaîne, entraîne un phénomène de gonflement de cette dernière. Elle occupe un volume apparent plus important.

Les détails des calculs sont contenus dans les ouvrages de références [De 79a, DJ87, Yam71]. Dans un premier temps, intéressons-nous à la description d'un **polymère isolé**, en l'absence d'interactions à longue distance.

## 2.2 MODÈLES FLEXIBLES

Dans une **chaîne idéale**, il n'existe aucune interaction entre monomères, ni entre la chaîne et le solvant. Nous considérons, ici, une **chaîne fantôme** c'est-à-dire qu'elle ne se voit pas elle-même. Ce modèle peut être considéré comme le point de départ de tous les autres modèles.

### 2.2.1 Modèle de la chaîne gaussienne sur un réseau

La chaîne idéale est constituée de  $N$  segments rigides de longueur  $l$ . L'orientation de ces segments est définie par les vecteurs  $\mathbf{r}_i$  avec  $i \in [1 : N]$ . Elle est **aléatoire** et **équiprobable** dans l'espace défini par le réseau à 3 dimensions  $(X, Y, Z)$ . Les vecteurs ne possèdent aucune autre restriction géométrique. Le chemin suivi par le polymère est une **marche aléatoire**, Fig. 2.1.

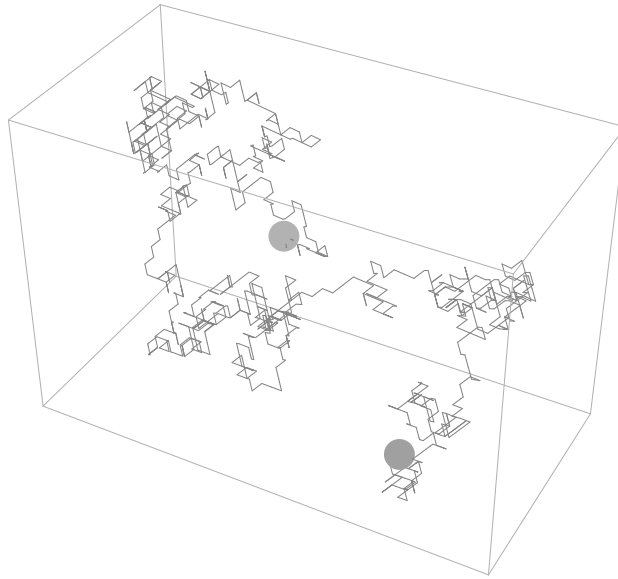


Fig. 2.1 – Chaîne Gaussienne,  $N = 1000$ , dans un espace 3D . Le point de départ est en rouge et le point d'arrivée en bleu

Sur notre objet ADN, ce modèle équivaudrait à ce qu'aucune des liaisons chimiques ne soit soumise à des contraintes angulaires. Il en résulte que les sous-parties de l'ADN sont considérées comme totalement **indépendantes**. Cette absence de contraintes est caractérisée comme un polymère totalement **flexible**. Autrement dit, ce modèle de polymère correspond à une chaîne ne comportant aucune interaction, ni à **courte distance**, ni à **longue distance**.

### Probabilité

À chaque **pas** de cette marche aléatoire, le vecteur  $\mathbf{r}_i$  peut prendre uniformément l'une des 6 directions de l'espace à 3 dimensions. La distribution de probabilité de la somme des vecteurs  $\mathbf{r}_i$  suit une loi gaussienne, si et seulement si  $N \gg l$ . Regardons l'expression de cette probabilité pour le cas d'une chaîne gaussienne dans un espace à 3 dimensions. La densité de probabilité,  $P$ , d'arriver au point  $\mathbf{R}$ , partant de l'origine, en  $N$  pas, s'écrit :

$$\begin{aligned} P(\mathbf{R}) &= P_x(\mathbf{R})P_y(\mathbf{R})P_z(\mathbf{R}) \\ &= Cte \cdot \exp\left[-\frac{x^2}{2\langle x^2 \rangle}\right] \cdot \exp\left[-\frac{y^2}{2\langle y^2 \rangle}\right] \cdot \exp\left[-\frac{z^2}{2\langle z^2 \rangle}\right] \end{aligned} \quad (2.1)$$

Ci-dessous, nous proposons des exemples de conformations, sur un réseau 2D, accessibles à la chaîne qui permettent de construire sa distribution de probabilité. La distribution de probabilité représentée en Fig. 2.2, correspond à la répartition de l'ensemble des mesures de la distance entre le point de départ et le point d'arrivée **selon l'axe x**.

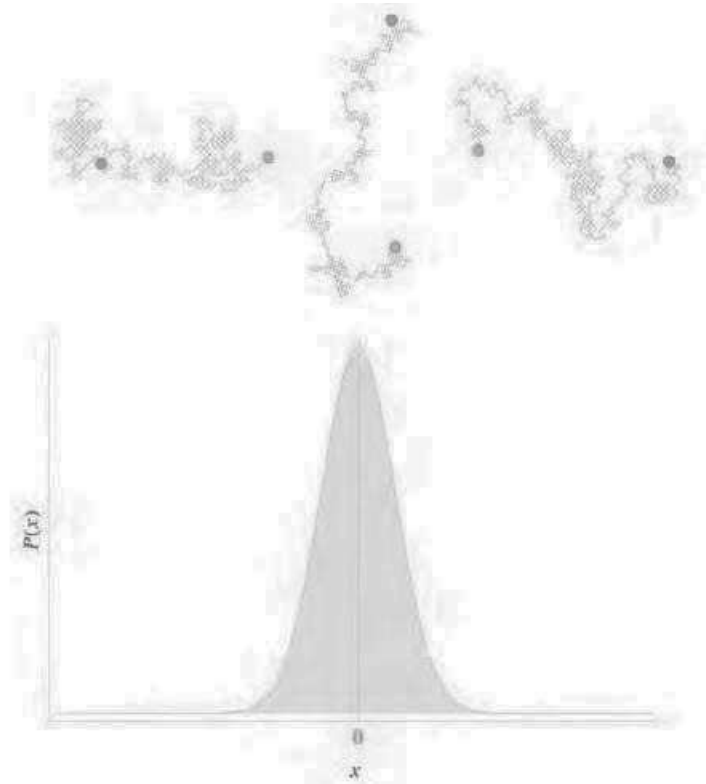


Fig. 2.2 – Distribution de probabilité de la chaîne selon  $x$ . Le choix de l'orientation de l'axe  $x$  des chaînes idéales, est en accord avec l'axe  $x$  du graphique. Le point de départ est en rouge et le point d'arrivée en bleu.

L'étalement de la distribution par rapport au centre de masse de la chaîne, suit une **loi d'échelle** :  $\Delta r^2 = Nl^2$ . En effet, l'étalement de la distribution augmente lorsque  $N$  augmente, illustration en Fig. 2.3.

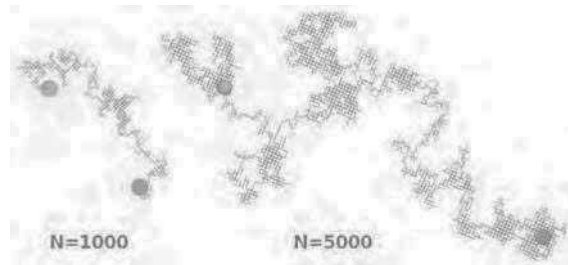


Fig. 2.3 – Représentation de chaînes idéales en 2 dimensions pour différentes valeurs de  $N$

La quantification de cet étalement peut se faire via la mesure de l'écart quadratique moyen  $\sigma^2$  directement accessible sur la distribution de probabilité. Ce terme correspond à **une mesure des fluctuations autour de la position moyenne**.

La construction de la distribution de probabilité dans l'ensemble des trois directions de l'espace ( $X, Y, Z$ ) conduit à la superposition des distributions  $P_x(\mathbf{R})$ ,  $P_y(\mathbf{R})$  et  $P_z(\mathbf{R})$ . La résolution de la densité de probabilité de présence,  $P(\mathbf{R})$ , en position finale  $\mathbf{r}_N$  après  $N$  pas de l'équation 2.1, est égale à :

$$P(\mathbf{R}) = \left( \frac{2}{3} \pi N l^2 \right)^{3/2} \exp \left[ -\frac{3}{2} \frac{|\mathbf{R}|^2}{N l^2} \right] \quad (2.2)$$

### 2.2.2 Modèle de la chaîne librement jointe : *Freely Jointed Chain (FJC)*

Le modèle de **Freely Jointed Chain** est construit sur le même principe qu'une **chaîne idéale**. Le modèle **FJC** correspond, lui aussi, à un modèle de polymère simplifié, Fig. 2.4.

Comme précédemment, la **chaîne librement jointe** est approximée par une série de  $N$  **segments statistiques** consécutifs de longueur  $l$ . Leurs directions sont déterminées par des vecteurs  $\mathbf{r}_i$  aléatoires dont les orientations sont totalement indépendantes les unes des autres, il n'y a plus de réseau sous-jacent. Une **chaîne librement jointe** est **gaussienne** dans la limite où  $N \rightarrow \infty$ .

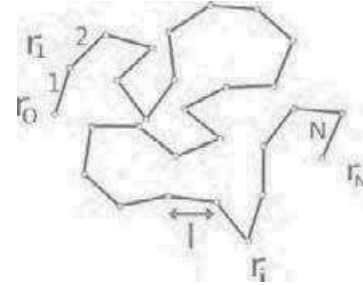


Fig. 2.4 – *Freely Jointed Chain*

Du fait de cette répétition d'un grand nombre d'unités élémentaires, le polymère est un objet statistique. Cette chaîne "*librement liée*" est parfaitement décrite par une **marche aléatoire**. Les segments sont considérés comme statistiquement indépendants, autrement dit l'orientation de ces derniers est décorrélée. Ici, cette décorrélation est valide au-delà de la longueur  $l$  du segment.

La conformation de la chaîne d'ADN ainsi décrite, à partir d'une marche aléatoire, est parfaitement analogue à la trajectoire d'une particule soumise au mouvement brownien. Le nombre de segments de la marche aléatoire joue le rôle du temps dans un mouvement brownien. Ainsi, la mesure de l'étalement de la gaussienne est une mesure de l'effet diffusif.

#### Longueur de contour : $L$ (en : *contour length*)

La **longueur de contour** correspond à la distance entre les deux extrémités de la chaîne si nous considérons celle-ci comme totalement étirée. La **longueur de contour** correspond à la somme de tous les **segments statistiques**.

$$L = Nl \quad \text{Longueur de contour} \quad (2.3)$$

#### Distance bout-à-bout : $R$ (en : *end-to-end distance*)

Une des informations qu'il nous est possible d'extraire de cette chaîne de polymère est la **distance bout-à-bout** que nous noterons  $R$ , Fig. 2.5.

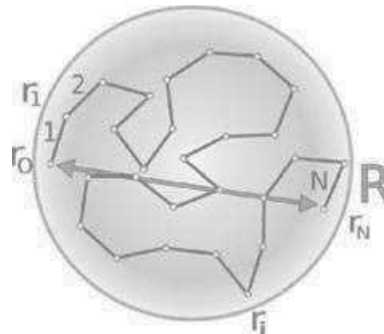


Fig. 2.5 – Distance bout-à-bout,  $R$ , dans un espace 2D. Nombre de segments  $N=27$

La **distance bout-à-bout** est la norme du vecteur  $\mathbf{R}$  qui s'exprime telle que :

$$\mathbf{R} = \mathbf{r}_N - \mathbf{r}_0 = \sum_{i=1}^N \mathbf{r}_i \quad (2.4)$$

Intéressons-nous à la **moyenne sur l'ensemble des conformations** de la chaîne. De part la définition même de la chaîne, qui est constituée de maillons **statistiquement indépendants**, l'orientation des vecteurs est aléatoire. De par cette isotropie, la **valeur moyenne de  $R$**  est nulle. En effet, pour toute conformation caractérisée par un vecteur  $\mathbf{r}_i$ , il existe une conformation symétrique dont le vecteur est  $-\mathbf{r}_i$ . Ainsi :

$$\begin{aligned} \langle \mathbf{R} \rangle &= \left\langle \sum_{i=1}^N \mathbf{r}_i \right\rangle = \sum_{i=1}^N \langle \mathbf{r}_i \rangle \\ &= 0 \end{aligned} \quad (2.5)$$

La grandeur caractéristique pertinente est donc l'**écart quadratique moyen entre les extrémités de la chaîne** :  $\langle \mathbf{R}^2 \rangle$ .

$$\begin{aligned} \langle \mathbf{R}^2 \rangle &= \left\langle \sum_{i=1}^N \sum_{j=1}^N \mathbf{r}_i \cdot \mathbf{r}_j \right\rangle = \sum_{i=1}^N \sum_{j=1}^N \langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle \\ &= \sum_{i=1}^N \langle \mathbf{r}_i \cdot \mathbf{r}_i \rangle + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle \end{aligned} \quad (2.6)$$

Dans le cas où l'orientation des segments est aléatoire et sans corrélation, le terme  $\langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle$  est nul pour tout  $i \neq j$ , il ne reste alors que les termes où  $i = j$ .

L'**écart quadratique moyen entre les extrémités de la chaîne** égale alors :

$$\langle \mathbf{R}^2 \rangle = \sum_{i=1}^N \langle \mathbf{r}_i^2 \rangle = Nl^2 \quad R_{FJC} \quad (2.7)$$

Ceci correspond au résultat classique d'une **marche aléatoire**.

La mesure de  $\langle \mathbf{R}^2 \rangle$  répond aussi à la question suivante : à quelle distance du centre de masse  $G$  se trouve typiquement le point d'arrivée  $r_N$  après  $N$  pas ?

### Rayon de giration : $R_g$ (en : *radius of gyration*)

Une autre longueur caractéristique de la chaîne est le **rayon de giration**  $R_g$ . Le **rayon de giration** correspond à la **valeur quadratique moyenne** de la taille de la chaîne. Cette longueur est une mesure de la distance moyenne entre le centre de masse de la chaîne,  $G$ , et les positions des monomères le long de la chaîne.

Soit  $\mathbf{R}_g$  le vecteur définissant la position du  $i$ -ème segment de la chaîne par rapport à son centre de masse, Fig. 2.6 :

$$\langle \mathbf{R}_g^2 \rangle = \frac{1}{2(N+1)^2} \left\langle \sum_{i=1}^N |\mathbf{r}_i - \mathbf{R}_g|^2 \right\rangle \quad (2.8)$$

$$= \frac{N(N+2)}{6(N+1)} l^2 \quad (2.9)$$

$$N \rightarrow \infty \Rightarrow \langle \mathbf{R}_g^2 \rangle = \frac{1}{6} N l^2 \quad (2.10)$$

Suite aux changements incessants de conformation, dus à l'agitation thermique, la chaîne prend, en moyenne, la forme d'une **pelote statistique**. L'ensemble de la chaîne est contenu dans une sphère de rayon typique  $R_g = \sqrt{\langle \mathbf{R}_g^2 \rangle}$ .

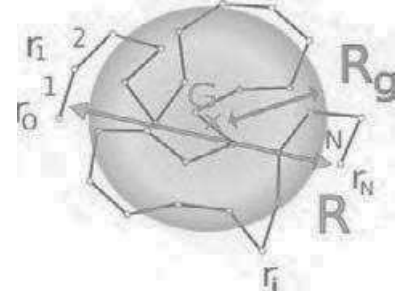


Fig. 2.6 – Rayon de giration,  $\mathbf{R}_g$ , dans un espace 2D. Nombre de segments  $N=27$

Toujours pour une longue chaîne où  $N \rightarrow \infty$ , le rapport du **rayon de giration** sur la **distance bout-à-bout** est indépendant de la structure de la chaîne. En effet, ce rapport est seulement égal à un facteur numérique, il ne dépend pas de  $l$ .

$$\frac{\langle \mathbf{R}_g^2 \rangle}{\langle \mathbf{R}^2 \rangle} = \frac{1}{6} \quad (2.11)$$

### L'entropie : $S$

Dans le modèle **FJC**, l'ensemble des configurations accessibles, ou micro-états, possède la même énergie. Par définition l'entropie,  $S$ , est donc ici, le produit de la constante de Boltzmann  $k_B$  par le logarithme du nombre de micro-états,  $P(\mathbf{R})$ , accessibles à la chaîne constituée de  $N$  monomères et de vecteurs bout-à-bout  $\mathbf{R}$ .

$$S = k_B \ln(P(\mathbf{R}, N)) \quad (2.12)$$

$$\mathbf{R} \ll L \Rightarrow S = k_B \frac{3 |\mathbf{R}|^2}{2 N l^2} + S(N, \mathbf{R} = 0) \quad (2.13)$$

Ceci correspond à la réponse entropique linéaire d'une chaîne polymère soumise à de faibles extensions.

Le modèle **FJC** permet, notamment, de calculer la longueur d'extension d'un polymère en fonction de la force d'étirement à laquelle il est soumis.

Rappelons que les conformations possibles de cette chaîne sont construites par des maillons **statistiquement indépendants** dont l'orientation est aléatoire et équiprobable. Ceci signifie que l'orientation du maillon  $\mathbf{r}_i$  n'influe aucunement sur l'orientation de maillon  $\mathbf{r}_{i+1}$  ou  $\mathbf{r}_{i-1}$ . Pourtant, soumise à l'énergie d'agitation thermique, l'orientation des segments statistiques est porteuse d'une certaine **élasticité** qui est de nature purement **entropique**.

Pour nous en convaincre, il suffit d'appliquer une force à chaque extrémité de la chaîne afin d'imposer l'alignement de tous les vecteurs  $\mathbf{r}_i$ . Ainsi, due à l'agitation thermique, la réponse du polymère tend à désordonner cette orientation forcée des segments. Cette réponse élastique du polymère est purement **entropique**.

La forme de **pelote statistique** maximise l'entropie de la chaîne [De 79b]. La chaîne se comporte comme un **ressort entropique**. L'étirement de la chaîne réduit le nombre de configurations accessibles, il en résulte une **force de rappel** qui tend à désordonner les segments, à cause de l'agitation thermique.

### 2.2.3 Limite des modèles flexibles pour décrire une chaîne réelle

En premier lieu, il est important de garder à l'esprit que les paramètres caractéristiques que nous extrayons de ces modèles statistiques sont **moyennés** sur l'ensemble des conformations. Ces modèles ne rendent pas compte d'une conformation **instantanée** de chaque chaîne **individuelle** mais correspondent à une **moyenne d'ensemble**.

Les chaînes réelles, tel l'ADN, ne sont pas des polymères totalement flexibles. Nous l'avons vu, il existe des corrélations d'orientation entre les bases complémentaires ainsi qu'entre les paires de bases successives. En utilisant le segment élémentaire comme unité de base, le modèle simple de la chaîne librement jointe ne suffit pas à rendre compte correctement de l'élasticité de la molécule. En effet, il ne prend pas en compte l'élasticité du polymère sur des petites longueurs, de l'ordre de la taille des segments de la chaîne.

L'idée sous-jacente développée lors du calcul de l'entropie de la chaîne, est que l'orientation de chaque monomère par rapport à ses voisins immédiats n'est pas toujours indépendante pour les polymères réels. Pour décrire convenablement cette élasticité, il faut donc utiliser un modèle physique qui introduit des corrélations entre les segments à courte distance. Il faut introduire une limitation de la flexibilité de la chaîne. La longueur statistique de Kuhn est le premier concept nous permettant de rendre compte de cela.

De plus, si nous cherchons à étudier une chaîne courte par rapport à cette longueur de Kuhn, nous pouvons être amenés à explorer un domaine où la statistique de la chaîne n'est pas encore gaussienne. Il existe ce même problème si la chaîne est très rigide. Il est donc nécessaire de disposer d'un modèle pouvant décrire correctement ces différentes situations.

#### Longueur statistique de Kuhn : $b$ (en : *Kuhn length*)

La **longueur statistique de Kuhn** [Kuh39],  $b$ , donne une idée de la rigidité de la chaîne. Cette longueur est une longueur effective de maillon, au-delà de laquelle les corrélations orientationnelles disparaissent entre monomères. La **longueur statistique de Kuhn** est égale à la distance nécessaire pour perdre ces corrélations d'orientation entre les segments. Nous verrons par la suite, que ce paramètre rejoint la **longueur de persistance** à un facteur numérique près.

Ainsi, il faut dissocier la longueur du segment  $l$  avec la longueur statistique de Kuhn,  $b$ .

Pour décrire une chaîne réelle par le modèle de la chaîne librement jointe, il faut remplacer la longueur du monomère par la longueur de Kuhn  $b$  et leur nombre par  $N_{eff} = L/b$ , Fig. 2.7. Nous obtenons une chaîne équivalente définie par  $L = Nl = N_{eff}b$ . Ainsi, la chaîne reste gaussienne pour  $N_{eff}$  suffisamment grand.

En d'autres termes, la longueur de Kuhn décrit la nature des interactions inter-segments du polymère. Ainsi, si l'orientation d'une partie de la chaîne influe sur l'orientation de la partie de la chaîne qui la suit et/ou précède, la rigidité du polymère sera plus importante que si ces deux parties possèdent des orientations totalement décorrélées. **La longueur statistique de Kuhn** décrit la longueur minimale entre deux points de la chaîne, telle que tout se passe comme s'ils étaient **décorrélés**.

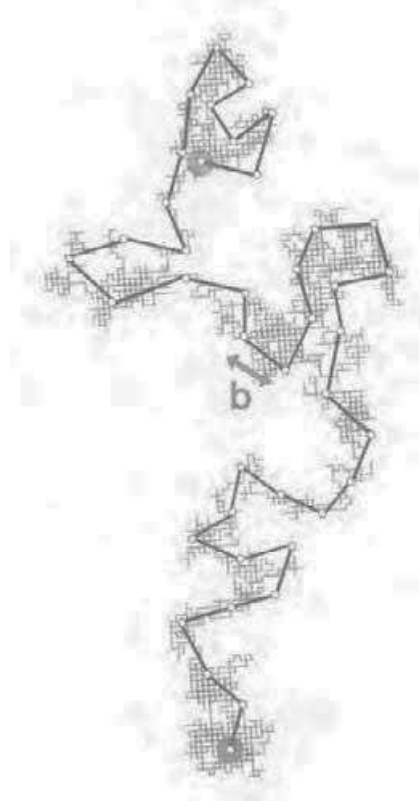
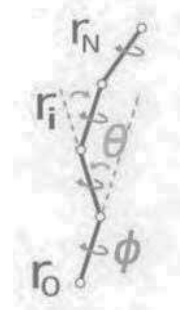


Fig. 2.7 – Représentation d'une chaîne idéale superposée avec sa chaîne équivalente de Kuhn

## 2.3 MODÈLES SEMI-FLEXIBLES

### 2.3.1 Modèle de Kratky-Porod (KP) ou *Freely Rotating Chain* (FRC)

Dans le modèle **Freely Rotating Chain** ou de **Kratky-Porod** la chaîne est de nouveau constituée de  $N$  segments successifs de longueur  $l$ , [KP49]. Dans ce modèle, les interactions à longue portée sont toujours absentes. Par contre, les interactions à courte portée sont prises en compte dans la construction même de la chaîne. En effet, chaque segment forme un angle  $\theta$  fixe avec le segment que le précède, tandis que l'azimut  $\phi$  est libre, d'où le nom du modèle, Fig. 2.8.



**Distance bout-à-bout**

Fig. 2.8 – *Freely Rotating Chain*

$$\langle R_{FRC}^2 \rangle = l^2 \left\langle \sum_{i=0}^N \cos^2 \theta \right\rangle \quad R_{FRC} \quad (2.14)$$

Lors du calcul de l'écart quadratique moyen sur la mesure de la distance bout-à-bout [De 79a], les termes croisés ne sont plus nuls comme dans l'Eq. 2.6. Le degré de liberté qui offre une certaine flexibilité à la chaîne se fait via la liberté de rotation selon l'angle  $\phi$ . L'originalité du modèle de **Kratky-Porod** réside dans l'introduction d'une courbure tout le long de la chaîne du polymère.



### Longueur de persistance : $L_p$ (en : *persistence lenght*)

Le modèle de **Kratky-Porod** permet de prendre en compte des corrélations d'orientations entre deux segments consécutifs  $\mathbf{r}_i$  et  $\mathbf{r}_{i+1}$ , contrairement au **modèle FJC** où  $\langle \mathbf{r}_i \cdot \mathbf{r}_{i+1} \rangle = 0$ . Cette corrélation entre maillons successifs induit des corrélations à plus longue distance sur la chaîne. Pour un tel modèle de chaîne, l'orientation entre deux segments  $\mathbf{r}_i$  et  $\mathbf{r}_j$  reste corrélée sur une distance nommée **longueur de persistance**,  $L_p$ . Au-delà de cette longueur, les corrélations orientationnelles sont, essentiellement, perdues. Elle est définie par :

$$\langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle = \exp\left(-\frac{|i-j|}{L_p}\right) \quad (2.15)$$

La **longueur de persistance intrinsèque**,  $L_p$ , est la longueur caractéristique sur laquelle la direction des segments reste corrélée. Au delà de cette longueur la chaîne a perdu l'essentiel de la mémoire de son orientation.

La longueur de persistance est mathématiquement définie par l'intermédiaire de la limite de la distance bout-à-bout lorsque le nombre de segments tend vers l'infini,  $N \rightarrow \infty$ . Dans cette limite, la valeur de la distance bout-à-bout est égale à  $\langle \mathbf{R}_{\text{FRC}}^2 \rangle = 2LL_p$ . La longueur de persistance approche alors la valeur :

$$L_p = \frac{l}{1 - \cos \theta} \quad (2.16)$$

À grande échelle, lorsque  $L \gg L_p$ , la longueur de persistance est une mesure de la rigidité intrinsèque de la chaîne du polymère.

#### 2.3.2 Modèle du "Ver" : *Worm Like Chain (WLC)*

Le modèle du ver est obtenu en faisant tendre  $l \rightarrow 0$ , soit une **limite continue** du modèle de Kratky-Porod, Fig. 2.9. Ceci revient à introduire une **courbure** tout le long de la chaîne.

La construction de ce modèle est donc la même que précédemment sauf que nous n'utilisons plus des segments de taille finie pour décrire la chaîne.

Dans ce modèle, la chaîne possède une courbure continue. Ceci introduit la notion de rigidité de courbure. En chacun de ses points, la courbure du polymère est décrite par la dérivée son vecteur tangent  $\mathbf{t}(s)$ , où  $s$  mesure l'abscisse curviligne.

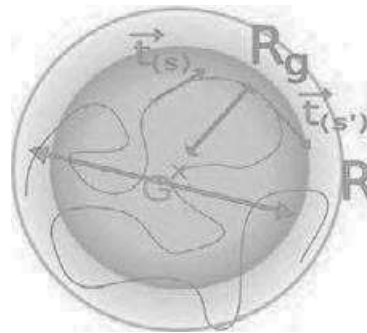


Fig. 2.9 – *Worm Like Chain*

### Longueur de persistance

La décorrélation des orientations des vecteurs tangents de la chaîne sur la **longueur de persistance intrinsèque**  $L_p$ , suit, de nouveau, une loi exponentielle :

$$\langle \mathbf{t}(s) \mathbf{t}(s') \rangle = \exp\left(-\frac{|s' - s|}{L_p}\right) \quad (2.17)$$

### Module de courbure

Modélisons notre objet comme une tige élastique. La capacité de ce dernier à se courber dépend de ses propriétés élastiques, c'est à dire de sa **longueur de persistance** et de son **module de courbure** :  $\kappa_c$ . Le module de courbure  $\kappa_c$  correspond à la multiplication du **module de Young**  $Y$  par le moment d'inertie d'une section du polymère,  $\mu_I$ . Il existe un lien étroit entre le module de courbure et  $L_p$

$$\Rightarrow L_p = \frac{\kappa_c}{k_B T} \quad (2.18)$$

La valeur numérique de la multiplication de la constante de Boltzmann  $k_B$  par la valeur de la température ambiante, soit  $T = 298$  K, est égale à  $\sim 4.10^{-21}$  J. Pour un ADN double brins le module de courbure est de l'ordre de  $\kappa_c \sim 7.10^{-29}$  J.m.

### Distance bout-à-bout et Longueur de persistance

La distance bout-à-bout  $\langle \mathbf{R}_{WLC}^2 \rangle$  de ce modèle se calcule via :

$$\langle \mathbf{R}_{WLC}^2 \rangle = \int_0^N \int_0^N \langle \mathbf{t}(s) \mathbf{t}(s') \rangle ds ds' \quad (2.19)$$

$$\langle \mathbf{R}_{WLC}^2 \rangle = \int_0^N \int_0^N \exp \left[ 1 - \frac{|s' - s|}{L_p} \right] ds ds' \quad (2.20)$$

Après deux intégrations successives, sur  $s'$  puis sur  $s$ , nous obtenons la distance bout-à-bout  $\langle \mathbf{R}_{WLC}^2 \rangle$  égale à :

$$\langle \mathbf{R}_{WLC}^2 \rangle = 2LL_p - 2L_p^2 \left[ 1 - \exp\left(-\frac{L}{L_p}\right) \right] \quad (2.21)$$

Lorsque la longueur du polymère est suffisamment grande,  $L \rightarrow \infty$  autrement dit  $L \gg L_p$ , le second terme de l'équation 2.21 s'annule. Nous obtenons donc une distance bout-à-bout telle que :

$$\langle \mathbf{R}_{WLC}^2 \rangle \simeq 2LL_p \quad (2.22)$$

Autrement dit nous retrouvons le résultat de l'équation 2.7 d'une chaîne à orientation libre suivant un **comportement gaussien**. Pour une telle chaîne, nous avons la définition suivante  $\langle \mathbf{R}_{WLC}^2 \rangle = Nl^2$ , qui ici est égale à  $\langle \mathbf{R}_{WLC}^2 \rangle = 2LL_p = N_{eff}b^2$ . Nous avons donc  $N_{eff} = \frac{L}{2L_p}$  et  $b^2 = (2L_p)^2$ . Ainsi, nous passons d'une chaîne où  $L = Nl$ , à une chaîne équivalente avec  $L = N_{eff}b$  et  $b = 2L_p$ . La longueur de Kuhn est donc égale à  $2L_p$ , dans la limite où le polymère est suffisamment long :  $L \rightarrow \infty$ .

*A contrario*, si  $L \ll L_p$  alors  $\langle \mathbf{R}_{WLC}^2 \rangle \simeq L^2$ . L'approximation de distance bout-à-bout du polymère est égale à la longueur de contour de la chaîne. La chaîne se comporte donc comme un objet sans aucune flexibilité : une **tige rigide** de longueur  $L$ .

Lorsque la longueur du polymère est du même ordre de grandeur que la longueur de persistance,  $L \sim L_p$ , alors nous sommes en régime **semi-flexible**, qui sera central dans cette thèse.

Lorsque nous nous intéressons à l'ADN, l'échelle de travail est le nanomètre :  $1.10^{-9} \text{ m} = 1 \text{ nm}$ . Dans la suite de cette étude, tous les paramètres de taille extraits de la caractérisation de nos ADN sont exprimés en [nm]. La longueur de persistance peut varier de quelques dixièmes de nanomètre à plusieurs centaines en fonction de la nature du système étudié, mais aussi en fonction de l'échelle d'observation. La valeur de la longueur de persistance classiquement renseignée pour l'ADN double brins est  $L_{p_{Ref}} \approx 50 \text{ nm}$  [Hag88]. Dès à présent, nous utilisons cette unité comme référence dans notre discours. Si nous nous intéressons à la fibre de chromatine, c'est-à-dire lorsque l'ADN double brins s'enroule périodiquement autour des histones composant les nucléosomes comme cela se produit dans le noyau cellulaire, alors la longueur de persistance de ce système serait différente de l'ADN double brin seul. Les niveaux d'organisations du biopolymère influent sur ses propriétés physiques et mécaniques, ainsi, la mesure de la longueur de persistance de la molécule d'ADN dépend de sa structure. En effet, une récente étude, [HMR<sup>+</sup>13], estime la longueur de persistance associée à l'ADN sous forme de chromatine, près de deux fois inférieure à  $L_{p_{Ref}} \approx 50 \text{ nm}$ , cette dernière reste bien entendu liée à celle de l'ADN double brin [BHLV01]. Pour comparaison avec la molécule d'ADN, une molécule d'actine, molécule constituant l'ossature rigide du cytosquelette des cellules eucaryotes, possède une longueur de persistance beaucoup plus importante :  $L_p \approx 15.10^3 \text{ nm}$  [YNNO84].

### Contribution des modèles semi-flexibles

L'existence de corrélation inter-segments contraint le domaine de conformation accessible à la chaîne. Ceci peut-être assimilé à des interactions répulsives qui tendent donc à augmenter la taille de la chaîne. Néanmoins, quel que soit le modèle utilisé dans les conditions où  $N \rightarrow \infty$ , la statistique de distribution reste gaussienne. Notamment la loi d'échelle, rendant compte de la proportionnalité entre  $\langle R^2 \rangle$  et  $N$ , reste vérifiée.

Il est intéressant de remarquer que les expressions qui viennent d'être données sont valables pour les polymères libres, non confinés, si bien que la forme des molécules n'est dictée que par la maximisation de leur entropie conformationnelle.

Ce modèle permet de rendre compte de la flexibilité et rigidité **intrinsèques** à la chaîne-même du polymère. Si  $L \ll L_p$  cette flexibilité n'est plus dépendante du nombre de segments  $N$  choisis pour décrire la chaîne. Ainsi, cela permet une étude pertinente d'un large spectre de longueurs du polymère, allant de séquences courtes et rigides, jusqu'à la forme en pelote.

#### 2.3.3 Limite des modèles semi-flexibles pour décrire une chaîne réelle

S. Smith et al. [SCB96] ont mesuré l'élasticité d'une seule molécule d'ADN double brins et simple brin. La molécule d'ADN est soumise à des forces d'étirement via une technique de pince optique. Leurs résultats expérimentaux décrivent la réponse de l'ADN, en terme de longueur d'extension, en fonction de la force exercée, Eq. 2.13. Seul le modèle du ver est en excellent accord avec ces résultats.

Les modèles **semi-flexibles** permettent de décrire convenablement l'ADN afin de caractériser sa réponse conformationnelle sous certaines contraintes et conditions. Seulement, ces modèles restent insuffisants pour rendre compte de la subtilité de comportement de l'ADN *in vivo*.

Ainsi, jusqu'à présent, les modèles présentés ne sont pas auto-évitants, le polymère est autorisé à se recouper. Or, le squelette de l'ADN présente des interactions répulsives, notamment de nature électro-statique, qui interdisent ce type de croisements. Ceci correspond aux contraintes dues à la **structure intrinsèque** de l'ADN.

De plus, aucun terme ne rend compte de l'influence de l'**environnement local** sur la structure et les propriétés du polymère. Tout ceci implique des restrictions sur les conformations accessibles à la chaîne. Ces paramètres correspondent à des interactions dites à longue portée car elles découlent de la rencontre d'une sous-partie de la chaîne avec elle-même. Initialement, ces deux sous-parties ne sont pas spatialement proches du fait de leur position le long du contour de la molécule, mais soumises aux fluctuations de conformation, elles interagissent.

Il est donc nécessaire de se munir d'outils qui nous permettront de décrire plus justement le comportement des chaînes réelles de polymères. Ce qui suit va nous aider à affiner les propriétés statistiques de la chaîne dues aux effets des interactions à longue distance entre les monomères ainsi que des effets liés au solvant.

## 2.4 CHAÎNE RÉELLE

Pour décrire les chaînes réelles, nous devons donc ajouter des interactions à longue distance. En effet *in vivo*, l'ADN est en solution, ce qui implique et influe sur trois types d'interactions.

- Les interactions polymère-solvant
- Les interactions solvant-solvant
- Les interactions polymère-polymère

Les interactions qui nous intéressent sont bien-sûr celles faisant intervenir notre objet d'intérêt qu'est le polymère. Nous nous restreindrons à leur caractérisation. De plus, les principales caractéristiques qui influent sur les propriétés du polymère sont la nature du solvant et la température du milieu.

La chaîne d'ADN occupe un certain volume. En fonction des conditions environnementales, ce dernier est plus ou moins important comme nous allons le voir. Ce volume apparent du polymère influe sur les différentes interactions citées plus haut. Nous allons voir que la qualité du solvant, bon ou mauvais, impacte sur l'effet de volume exclu et de gonflement de la chaîne d'ADN.

Dans cette étude le système utilisé est un double brin d'ADN s'approchant de la forme d'hélice B, sauf mention expresse du contraire.

### 2.4.1 Théorie de Flory

#### Effet de volume exclu

L'effet de **volume exclu** correspond à l'impossibilité pour deux monomères d'occuper la même position dans l'espace. Le *volume* occupé par une partie de la chaîne est *exclu* pour l'autre. Jusqu'à présent nous avons seulement considéré une chaîne **fantôme**, invisible pour elle-même. Maintenant nous considérons chaque monomère comme un objet dur, impénétrable, autrement dit la chaîne (ou marche) est auto-évitante.

La chaîne ainsi considérée possède une taille moyenne plus étendue que dans les descriptions précédentes. L'effet de **volume exclu** autour de chaque monomère induit un effet de **gonflement de la chaîne**. Effectivement, ce volume devenu inaccessible et impénétrable opère comme des interactions répulsives entre la chaîne et elle-même. L'espace occupé par la chaîne en est donc agrandi. De plus, cet effet de volume exclu ajoute des contraintes qui tendent à diminuer la flexibilité de la chaîne, autrement dit à augmenter sa rigidité. La longueur de persistance que nous avons jusqu'alors considérée correspond à une propriété locale. Ici, ce phénomène induit une rigidité provenant des interactions entre segments éloignés de la chaîne, lorsqu'ils viennent se rencontrer.

### Gonflement de la chaîne

Pour quantifier l'effet de gonflement de la chaîne induit par ce critère d'auto-évitement, intéressons-nous à la mesure de  $\langle \mathbf{R}_{WLC}^2 \rangle$ . Reprenons le résultat de la distance bout-à-bout obtenue pour une chaîne *librement jointe*, Eq. 2.7 .

$$\langle \mathbf{R}_{FJC}^2 \rangle = Nl^2 \quad (2.23)$$

Si maintenant nous appliquons les propriétés d'une marche auto-évitante, il apparaît un exposant,  $\nu$  dit **exposant de Flory**

$$\langle \mathbf{R}_{FJC,F}^2 \rangle = N^{2\nu} l^2 > Nl^2 \quad \text{Théorie de Flory} \quad (2.24)$$

La première estimation de cet exposant, par Flory, donne  $\nu = 3/5$ . Des études ont calculé  $\nu$  par des méthodes de renormalisation, afin d'affiner sa valeur numérique à 0.588 [BLZJ76]. La première estimation de Flory reste extrêmement satisfaisante.

L'effet de **volume exclu** réduit notablement le nombre de conformations accessibles à la chaîne.

Comme nous l'avons vu précédemment, l'étalement de la chaîne est d'autant plus important que la chaîne est grande :  $N \rightarrow \infty$ . Plus une chaîne est grande plus la probabilité de rencontre de la chaîne avec elle-même augmente. Il en résulte que plus la chaîne est grande, plus les restrictions liées à l'effet de **volume exclu** sont sensibles. Sur des chaînes courtes ce phénomène existe toujours. Seul le taux de rencontre de la chaîne avec elle-même diminue, il en découle donc une diminution de l'effet de volume exclu.

### Cas de la molécule d'ADN

La cause physique du **volume exclu** est intimement liée à la nature du polymère et du solvant. En effet, nous avons vu précédemment que le squelette de l'ADN porte deux charges négatives par paire de bases. Ces charges, dirigées vers l'extérieur de la molécule dans les formes A et B de l'ADN, sont les premières à *intervenir* lorsque deux segments de la chaîne se rencontrent. Il existe alors des interactions répulsives entre les charges négatives des groupements phosphates. Cette propriété chimique contribue fortement au phénomène de volume exclu.

De plus, la nature même du solvant influe sur ce phénomène d'exclusion. En effet, l'*amplitude* de l'effet de volume exclu dépend de la nature et des propriétés

chimiques du solvant. Les molécules qui constituent le solvant viennent interagir et s'insérer avec les groupements de l'ADN. Les charges positives présentes en solution viennent interagir avec les charges négatives portées par l'ADN, diminuant ainsi l'effet de volume exclu.

Si l'effet de volume exclu est quantifiable cela signifie que les répulsions polymère-polymère l'emportent sur les interactions polymère-solvant. Dans ce cas, le polymère est correctement décrit par une chaîne auto-évitante. L'étude des polymères a permis de mettre en exergue l'existence de différentes qualités de solvant. Nous allons répertorier leurs principales caractéristiques.

### Bon et mauvais solvant

Un **bon solvant**, ou solvant athermal, est caractérisé par l'exposant de Flory décrit plus haut. La chaîne est soumise à un effet de volume exclu ainsi que de gonflement. Les monomères sont uniformément et homogènement repartis dans le volume occupé par la chaîne. La distance bout-à-bout suit la théorie de Flory définie en Eq. 2.24. Dans cette situation, les interactions à longue distance polymère-polymère sont de nature répulsive.

Dans les **solvants  $\theta$** , la chaîne adopte un rayon de giration et une distance bout-à-bout idéaux, comme décrit en Eq. 2.9 et Eq. 2.7. La gamme de valeur de l'exposant de Flory pour les **solvants  $\theta$**  est d'autant plus étroite que la taille du polymère est grande. Dans ce cas, les interactions répulsives et attractives à longue distance polymère-polymère se compensent.

Enfin, dans les **mauvais solvants**, la chaîne s'effondre sur elle-même, elle se condense. Le rayon de giration devient alors inférieur à la valeur idéale donnée en Eq. 2.9. La chaîne adopte une conformation dense dite de type *globule*. Les interactions à longue distance polymère-polymère sont alors attractives.

L'étude de l'exposant de Flory a fait apparaître la dépendance de ce dernier en fonction de la température du milieu. Nous pouvons donc résumer les trois cas définissant la qualité du solvant de la façon suivante :

- Bon solvant  $\Rightarrow \nu > \frac{1}{2}$  &  $T > T_{critique}$
- Solvant  $\theta \Rightarrow \nu = \frac{1}{2}$  &  $T = T_{critique}$
- Mauvais solvant  $\Rightarrow \nu < \frac{1}{2}$  &  $T < T_{critique}$

Tous ces raffinements à ajouter au modèle de polymère impliquent une perte d'entropie de la chaîne car ils restreignent les degrés de liberté accessibles. De plus, la statistique gaussienne précédemment établie est mise à mal. L'existence de ces interactions entre segments non voisins, en bon solvant, a pour effet de gonfler la chaîne par rapport à ses conditions **idéales**. Cela revient à considérer deux forces en présence : une force de gonflement générée par les répulsions entre segments (effets de volume exclu) à laquelle vient s'opposer une force de rappel élastique relative à la diminution d'entropie de la chaîne déformée par rapport à son état idéal.

## CONCLUSION DU CHAPITRE

L'apport des modèles de physique statistique à la compréhension des objets du vivant est importante. En effet, ces modèles décrivent des systèmes simples permettant une étude des conformations et propriétés de la molécule d'ADN. Les informations extraites rendent compte des caractéristiques statistiques moyennes des conformations accessibles à la chaîne. Ces modèles sont construits sur différentes échelles de complexité et de discrétisation. Ils prennent en compte la nature entropique de l'élasticité du polymère. De plus, il est possible d'y inclure les interactions existant à courte ainsi qu'à longue distance.

Nous verrons dans le chapitre suivant que la longueur caractéristique de la chaîne qui nous intéresse est  $\sqrt{\langle \mathbf{R}^2 \rangle}$ , la valeur moyenne de la distance bout-à-bout de la molécule d'ADN. Par la suite nous noterons cette grandeur  $R$ . En nous focalisant sur cette grandeur, soulignons les aspects importants à retenir pour une bonne utilisation des différents modèles que nous venons de décrire.

Tout d'abord le choix du modèle est primordial. En effet, le calcul de la distance bout-à-bout pour une chaîne **librement jointe** ou une chaîne décrite par le **modèle du ver** n'amène pas au même résultat. Considérons un ADN de longueur de contour  $L = 200$  et  $2000$  nm. Nous exprimons le modèle **FJC**, Eq. 2.7, en prenant en compte des corrélations sur la longueur de Kuhn Eq. 2.22. Autrement dit, l'ADN correspond à une chaîne constituée de  $N = L/l$  maillons de longueur  $l = 0.34$  nm. Tandis que l'expression du modèle **WLC** suit l'Eq. 2.21. Ceci conduit aux expressions des modèles **FJC** et **WLC** suivantes :

$$\begin{aligned} - \langle \mathbf{R}_{\text{FJC}}^2 \rangle &= 2LL_p \\ - \langle \mathbf{R}_{\text{WLC}}^2 \rangle &= 2LL_p - 2L_p^2 (1 - \exp(-\frac{L}{L_p})) \end{aligned}$$

Sondons une large gamme de longueurs de persistance :  $L_p \in [0 : 100]$  nm. Que devient alors la valeur de  $R$  ?

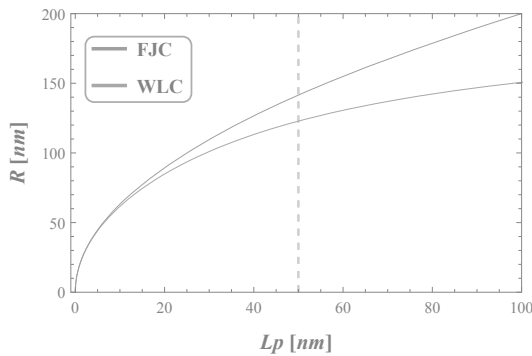


Fig. 2.10 – Comparaison des modèles **FJC** et **WLC**. Représentation de l'influence du modèle sur la distance bout-à-bout avec  $L = 200$  nm

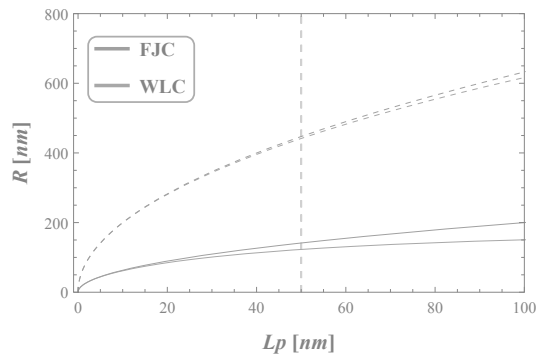


Fig. 2.11 – Comparaison des modèles **FJC** et **WLC**, de l'influence de  $L$  sur  $R$ . Courbes pleines  $L = 200$  nm, courbes pointillées  $L = 2000$  nm

L'écart entre ces deux modèles est évident. Pour  $L = 200$  nm, traits pleins Fig. 2.10 et 2.11, la variation de la valeur de  $R$  est de l'ordre de  $\sim 40$  nm autour de  $L_p = 50$  nm, ligne pointillée grise verticale. Cet écart provient de la prise en compte des corrélations d'orientation des positions le long de la chaîne contenues dans le terme  $L_p$ , le modèle **FJC** ne prend pas en compte ce type de corrélation.

Cette comparaison souligne bien la diminution des conformations accessibles et donc la perte d'entropie de la chaîne décrite dans le modèle **WLC** et donc la diminution de  $R$ . Par contre, cet écart diminue fortement lorsque  $L = 2000$  nm, traits pointillés Fig. 2.11,  $\Delta R < 20$  nm. En effet, pour cette longueur d'ADN, la chaîne se rapproche de la description gaussienne  $L \sim 20.b$ . Les deux modèles tendent donc vers la même limite de **chaîne idéale**.

Si nous incrémentons le modèle en tenant compte des interactions à longue distance, nous devons donc ajouter des interactions répulsives entre maillons de la chaîne. Ce phénomène de volume exclu interdit tout croisement de la chaîne avec elle-même et diminue d'autant plus l'entropie de la chaîne. Là encore, intéressons-nous aux différences induites par le choix du modèle : **FJC** et **WLC** ou la **théorie de Flory**. L'expression de la distance bout-à-bout selon cette dernière s'exprime via Eq. 2.24 où  $\nu = 3/5$ , à laquelle nous ajoutons la prise en compte des corrélations sur la longueur de Kuhn :

$$— \langle \mathbf{R}_{\text{FJC},F}^2 \rangle = 4L_p^2 \cdot \left( \frac{L}{2L_p} \right)^{6/5}$$

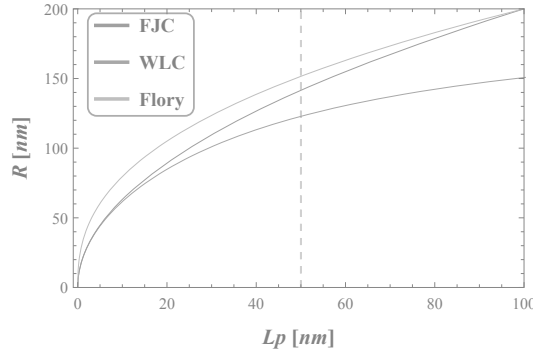


Fig. 2.12 – Comparaison des modèles **FJC** et **WLC** et **théorie de Flory**. Influence du modèle pour  $L = 200$  nm.

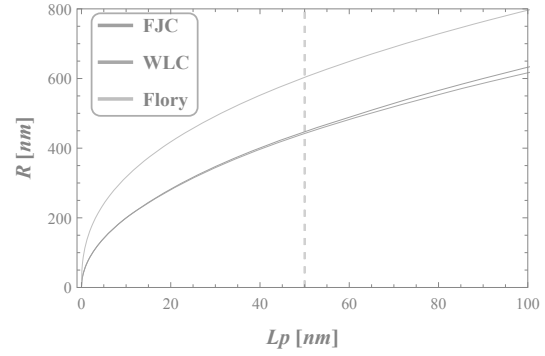


Fig. 2.13 – Comparaison des modèles **FJC** et **WLC** et **théorie de Flory**. Influence du modèle pour  $L = 2000$  nm.

L'effet de gonflement de la chaîne dû au volume exclu reste faible pour les polymères de petite taille car il y a peu de contacts, Fig. 2.12, par contre il devient d'autant plus important que l'ADN est long, Fig. 2.13. Ceci est en parfait accord avec l'augmentation de la probabilité de rencontre de la chaîne avec elle-même quand la longueur de cette dernière augmente.

L'étude de telles macromolécules converge vers un comportement général décrit dans les modèles présentés ci-dessus. En effet, si l'objet est assez long, les propriétés à grande échelle jouent un rôle aussi important sur les contraintes et les potentiels locaux associées aux propriétés chimiques et biologiques de la molécule d'ADN. Le comportement de l'ADN est alors décrit par ses paramètres effectifs. Le choix du modèle est donc crucial pour rendre pleinement compte des propriétés du système. La caractérisation et la description de phénomènes et contraintes **locaux** reste donc un enjeu. De plus, nous avons pu constater que certains phénomènes dépendent de la taille de la molécule d'ADN étudiée. Il est donc intéressant d'approfondir nos connaissances dans ces tailles d'ADN intermédiaires. En effet, c'est précisément cette gamme de longueur de molécule que nous allons étudier.



# CONCLUSION DE LA PARTIE

Comme nous venons de le voir, les aspects structuraux de la **molécule d'ADN** tels que les conformations des constituants moléculaires ou bien encore les angles de rotation, interviennent dans la conformation de la chaîne. L'existence de ces modulations structurales induit une augmentation des conformations accessibles et donc des degrés de libertés géométriques du système. Tout ceci confère une certaine flexibilité structurale à la molécule d'ADN. De plus, même après la prise en compte des contraintes induites par les corrélations conformationnelles entre ces constituants, la flexibilité structurale reste notable. Cela nous amène à considérer la double hélice d'ADN comme une structure qui ne possède pas un caractère totalement statique, dans l'espace et le temps. La molécule d'ADN est donc indissociable des notions de **flexibilité** et de **rigidité**.

Afin de caractériser les propriétés conformationnelles de la molécule d'ADN, nous utilisons des outils adaptés de physique statistique qui nous permettent de définir le **polymère ADN**. Il existe différents degrés de complexité de la chaîne. Une marche aléatoire idéale permet de décrire simplement une chaîne parfaitement **flexible**. Pour rendre compte de corrélations des positions de la chaîne sur des distances intermédiaires, de l'ordre de la longueur de persistance, nous utiliserons le modèle du polymère **semi-flexible**. L'utilisation de ces modèles permet de rendre compte de la courbure existant tout le long de la chaîne. Cette dernière est ainsi essentiellement caractérisée par ses propriétés physiques de **flexibilité** et de **rigidité** qui sont incluses dans la grandeur  $L_p$ .

La **molécule d'ADN** n'est jamais isolée. L'existence des différentes formes de double hélice est directement corrélée à l'**environnement** direct de l'ADN. *In vivo*, les molécules du solvant interagissent avec l'ADN et modifient ses propriétés physiques. Les **modèles de polymère** décrivant l'ADN sont intrinsèquement dépendants des **propriétés globales** telles que la flexibilité et la rigidité de l'ADN. Il semble donc facile de modifier et/ou d'implémenter ces modèles de physique statistique afin de rendre compte des variations de ces paramètres du fait de contraintes globales appliquées à la chaîne.

Une **contrainte globale**, telle qu'une variation de la concentration en sel du solvant, influe sur la conformation et les propriétés physiques de l'ADN. Quel impact peut avoir une variation du milieu sur les propriétés physiques, mécaniques et même structurales de l'ADN? Avons-nous les clés pour décrire et caractériser ce type de phénomène sur un modèle de polymère?

La structure de la **molécule d'ADN** est directement influencée par des **effets de séquence**. Nous l'avons vu, des séquences courtes riches en bases **C-G** peuvent fortement modifier l'agencement de la molécule jusqu'à la transition vers la forme **Z** de l'ADN. Là encore, les caractéristiques physiques et mécaniques de l'ADN sont influencées par ses **propriétés locales**. Les **modèles de polymère** décrivent la chaîne selon des paramètres de flexibilité et rigidité définis à l'échelle caractéristique de la longueur  $L_p$ . Une **modification locale** des propriétés de la chaîne peut-être modélisée par l'ajout de **contraintes locales** dans la description du polymère.

Un **défaut local**, telles qu'une courbure ou une bulle de dénaturation, peut influencer sur la géométrie et l'agencement conformationnel de l'ADN. Quel impact peut avoir une séquence bien spécifique sur les propriétés physiques, structurales et mécaniques de l'ADN ? Jusqu'à quelle échelle un effet local peut-il avoir une signature sur le comportement global de l'ADN ? Quel modèle de polymère rend compte au mieux, ce type de défaut ? Quelles sont les limites de ces modèles et leurs améliorations possibles ?

Afin de décrire l'existence et l'effet de corrélation sur des sous-parties de l'ADN pouvant être distantes d'une longueur supérieure à la longueur de persistance, les outils des *chaînes réelles* sont indispensables. Ces effets sont directement dépendants de l'échelle de notre système. En effet, comme nous l'avons vu, l'effet de volume exclu influe fortement sur la conformation d'une longue chaîne tandis qu'il est presque indiscernable sur le comportement de chaînes courtes,  $L$  de l'ordre de quelques  $L_p$ . Il est donc nécessaire et primordial d'adapter la description du polymère à l'échelle du système étudié.

Quelque soit le niveau des contraintes, local ou global, le défi réside en plusieurs points. Tout d'abord quelque soit l'effet étudié, variations des propriétés du milieu, courbure locale de l'ADN ou bien ouverture de la double hélice, il est nécessaire de caractériser son influence. Pour cela sonder la réponse de molécules d'ADN **individuelles** est préconisé. L'effet doit donc être décorrélé, le plus possible, de tous autres effets pouvant influencer sur la conformation de l'ADN. Il est nécessaire de caractériser son **intensité** avec précision ainsi que son amplitude en terme de *distance d'impact* ou d'**échelle d'influence**. Ensuite, l'**identification** des propriétés physiques modifiées par l'application de nos contraintes est primordial. Ceci nous permettra de **traduire** au mieux le comportement de l'ADN via les modèles de physique statistique. L'existence de **corrélations** entre différents paramètres, bien que délicate, doit être correctement définie. L'utilisation de différentes tailles de molécules d'ADN peut nous permettre de sonder **différents régimes** de réponse élastique. De même, prospecter différentes magnitudes de contraintes peut nous permettre de souligner les transitions de régimes. En dernier lieu, la réponse de l'ADN peut être mise en relation avec son rôle lors de **processus biologiques**, afin d'amener de nouveaux éléments à sa compréhension.

Dans notre étude nous avons donc choisi de **coupler** une approche **expérimentale** de **suivi de molécule unique** à des **simulations** basées sur un **algorithme de Monte-Carlo**.

## BIBLIOGRAPHIE

- [AE97] R. K. Allemann and M. Egli. DNA recognition and bending. *Chemistry & Biology*, 4(9) :643–650, September 1997.
- [AS72] C. Altona and M. Sundaralingam. Conformational analysis of the sugar ring in nucleosides and nucleotides. New description using the concept of pseudorotation. *Journal of the American Chemical Society*, 94(23) :8205–8212, November 1972.
- [Ast39] W T Astbury. X-Ray Studies of the Structure of Compounds of Biological Interest. *Annual Review of Biochemistry*, 8(1) :113–133, June 1939.
- [Ave44] O. T. Avery. Studies on the chemical nature of the substance inducing transformation of pneumococcal types : induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *Journal of Experimental Medicine*, 79(2) :137–158, February 1944.
- [BCT00] V A Bloomfield, D M Crothers, and I Tinoco Jr. *Nucleic Acids : Structures, Properties, and Functions*. University Science Books, illustrée edition, 2000.
- [BHLV01] E. Ben-Haïm, A. Lesne, and J-M. Victor. Chromatin : A tunable spring at work inside chromosomes. *Physical Review E*, 64, 2001.
- [BLZJ76] E Brezin, J C Le Guillou, and J Zinn-Justin. Phase Transitions and Critical Phenomena, Vol. VI. Eds C Domb, MS Green, New York : Academic, 1976.
- [CD97] C. R. Calladine and H. Drew. *Understanding DNA : The Molecule and How It Works*. 1997.
- [Cha50] E. Chargaff. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia*, 6(6) :201–209, June 1950.
- [Dau93] M. Daune. *Biophysique moléculaire : structures en mouvement*. Interediti edition, 1993.
- [De 79a] P-G. De Gennes. *Scaling Concepts in Polymer Physics*. 1979.
- [De 79b] P-G. De Gennes. *Scaling concepts in polymer physics*. Cornell university press, 1979.
- [Dic89] R. E. Dickerson. Definitions and nomenclature of nucleic acid structure parameters. *The EMBO journal*, 8(1) :1–4, January 1989.
- [Dic92] R. E Dickerson. DNA structure from A to Z. *Methods in enzymology*, 211 :67–111, 1992.
- [Dic99] R. E Dickerson. Helix structure and molecular recognition by B-DNA. *Nucleic Acid Structure*, 1, 1999.
- [Dio89] M Dion. Biologie terminale D, collection ADN. Hachette Lycées, 1989.
- [DJ87] J. Des Cloizeaux and G. Jannink. *Les polymères en solution : leur modélisation et leur structure*. Les éditions de physique, 1987.

- [DTT<sup>+</sup>80] H. Drew, T. Takano, S. Tanaka, K. Itakura, and R. E. Dickerson. High-salt d(CpGpCpG), a left-handed Z' DNA double helix. *Nature*, 286(5773) :567–573, August 1980.
- [DWT<sup>+</sup>81] H R Drew, R M Wing, T Takano, C Broka, S Tanaka, K Itakura, and R E Dickerson. Structure of a B-DNA dodecamer : conformation and dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 78(4) :2179–83, April 1981.
- [FANF11] A. Fiorini, Q. Alves, De.L. Neto, and M. A. Fernandez. Sequence-Directed DNA Curvature in Replication Origins Segments. In Jelena Kusic-Tisma, editor, *Fundamental Aspects of DNA Replication*, chapter 9. InTech, 2011.
- [FG53a] RE Franklin and RG Gosling. Evidence for 2-Chain Helix in Crystalline Structure of Sodium Deoxyribonucleate. *Nature*, 172 :156 – 157, 1953.
- [FG53b] RE Franklin and RG Gosling. Molecular Configuration in Sodium Thymonucleate. *Nature*, 171(4356) :740–741, April 1953.
- [Hag88] P J Hagerman. Flexibility of DNA. *Annual review of biophysics and biophysical chemistry*, 17 :265–286, 1988.
- [Hag90] P. J. Hagerman. Sequence-Directed Curvature of DNA. *Annual Review of Biochemistry*, 59(1) :755–781, June 1990.
- [Her52] A. D. Hershey. Independent functions of viral protein and nucleic acid in growth of Bacteriophage. *The Journal of General Physiology*, 36(1) :39–56, September 1952.
- [Her93] P. Herbomel. *L'expression du génome : du noyau à l'organisme*. De Boeck Secundair, 1993.
- [HMR<sup>+</sup>13] H. Hajjoul, J. Mathon, H. Ranchon, I. Goiffon, J. Mozziconacci, B. Albert, P. Carrivain, J-M. Victor, O. Gadal, K. Bystricky, and A. Bancaud. High-throughput chromatin motion tracking in living yeast reveals the flexibility of the fiber throughout the genome. *Genome research*, 23(11) :1829–38, November 2013.
- [HR96] A Herbert and A Rich. The biology of left-handed Z-DNA. *The Journal of biological chemistry*, 271(20) :11595–8, May 1996.
- [Hun93] C A Hunter. Sequence-dependent DNA structure. The role of base stacking interactions. *Journal of molecular biology*, 230(3) :1025–54, April 1993.
- [KP49] O. Kratky and G. Porod. Röntgenuntersuchung gelöster Fadenmoleküle. *Recueil des Travaux Chimiques des Pays-Bas*, 68(12) :1106–1122, September 1949.
- [Kuh39] W. Kuhn. Beziehungen zwischen Molekülgrösse, statistischer Molekülgestalt und elastischen Eigenschaften hochpolymerer Stoffe. *Kolloid-Zeitschrift*, vol. 87(1) :p. 3–12, 1939.
- [MA78] L. Michael and W. Arie. Extreme Conformational Flexibility of the Furanose d Ring in DNA and RNA. *Journal of the American Chemical Society*, 100 :2607–2613, 1978.

- [MS58] M. Meselson and F. W. Stahl. The replication of DNA in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 44(7) :671–682, July 1958.
- [MSWH58] D. A. Marvin, M. Spencer, M. H. F. Wilkins, and L. D. Hamilton. A New Configuration of Deoxyribonucleic Acid. *Nature*, 182(4632) :387–388, August 1958.
- [NCK<sup>+</sup>93] E. Nordhoff, R. Cramer, M. Karas, F. Hillenkamp, F. Kirpekar, K. Kristiansen, and P. Roepstorff. Ion stability of nucleic acids in infrared matrix-assisted laser desorption/ionization mass spectrometry. *Nucleic Acids Research*, 21(15) :3347–3357, July 1993.
- [PC53] L. Pauling and R. B. Corey. A Proposed Structure For The Nucleic Acids. *Proceedings of the National Academy of Sciences of the United States of America*, 39(2) :84–97, February 1953.
- [RZ03] A. Rich and S. Zhang. The structure of DNA in the nucleosome core. *Nature*, 423(6936) :145–50, May 2003.
- [SAB<sup>+</sup>96] T. R. Strick, J.-F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette. The Elasticity of a Single Supercoiled DNA Molecule. *Science*, 271(5257) :1835–1837, March 1996.
- [SCB96] S. B. Smith, Y. Cui, and C. Bustamante. Overstretching B-DNA : The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules. *Science*, 271(5250) :795–799, February 1996.
- [Sch14] H. Schiessel. *Biophysics for Beginners : a Journey through the Cell Nucleus*. January 2014.
- [SJHo4] J. Sponer, P. Jurecka, and P. Hobza. Accurate interaction energies of hydrogen-bonded nucleic acid base pairs. *Journal of the American Chemical Society*, 126(32) :10142–51, August 2004.
- [SJM<sup>+</sup>06] J. Sponer, P. Jurecka, I. Marchan, F. J. Luque, M. Orozco, and P. Hobza. Nature of base stacking : reference quantum-chemical stacking energies in ten unique B-DNA base-pair steps. *Chemistry (Weinheim an der Bergstrasse, Germany)*, 12(10) :2854–65, March 2006.
- [SRR76] N. C. Seeman, J. M. Rosenberg, and A. Rich. Sequence-specific recognition of double helical nucleic acids by proteins. *Proceedings of the National Academy of Sciences*, 73(3) :804–808, March 1976.
- [TDB<sup>+</sup>12] N. A. Temiz, D. E. Donohue, A. Bacolla, B. T. Luke, and J. R. Collins. The role of methylation in the intrinsic dynamics of B- and Z-DNA. *PloS one*, 7(4) :e35558, January 2012.
- [TV] L. Trabuco and E. Villa. CaseStudy : DNA.
- [VHHo1] J. M. Vargason, K. Henderson, and P. S. Ho. A crystallographic map of the transition from B-DNA to A-DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 98(13) :7265–7270, 2001.
- [WC53] J. D. Watson and F. H. C. Crick. Molecular Structure of Nucleic Acids : A Structure for Deoxyribose Nucleic Acid. *Nature*, 171(4356) :737–738, April 1953.

- [WDT<sup>+</sup>80] R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R. E. Dickerson. Crystal structure analysis of a complete turn of B-DNA. *Nature*, 287(5784) :755–758, October 1980.
- [WQK<sup>+</sup>79] A. H.-J. Wang, G.J. Quigley, F.J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel, and A. Rich. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature*, 282(5740) :680–686, December 1979.
- [WSW53] M H F Wilkins, a R Stokes, and H R Wilson. Molecular structure of deoxypentose nucleic acids., 1953.
- [Yam71] H. Yamakawa. *Modern Theory of Polymer Solutions*, 1971.
- [YNNO84] T. Yanagida, M. Nakase, K. Nishiyama, and F. Oosawa. Direct observation of motion of single F-actin filaments in the presence of myosin. *Nature*, 307(5946) :58–60, January 1984.

## Deuxième partie

# Notre approche à l'échelle de la molécule unique d'ADN





# TECHNIQUE EXPÉRIMENTALE : *Tethered Particle Motion (TPM)*

## SOMMAIRE

INTRODUCTION . . . . .	61
3.1 PRINCIPE DU <i>Tethered Particle Motion (TPM)</i> . . . . .	63
3.1.1 Préparation du support de l'échantillon . . . . .	64
Préparation de la lamelle de verre . . . . .	64
Dépôt de protéines . . . . .	64
3.1.2 Préparation de la chambre enfermant l'échantillon . . . . .	65
Préparation de la lame de verre . . . . .	65
Assemblage de la chambre . . . . .	66
3.1.3 Préparation des complexes ADN/particule . . . . .	66
ADN . . . . .	66
Coefficient de diffusion . . . . .	67
Nanoparticules . . . . .	67
Formation des complexes ADN/particule . . . . .	68
3.1.4 Finalisation de l'échantillon . . . . .	68
Diffusion de nos complexes ADN/particule dans la chambre microfluidique . . . . .	68
3.2 TPM-ON-A-CHIP . . . . .	68
3.3 POINTS FORTS DU TPM-ON-A-CHIP . . . . .	69
3.4 DISPOSITIF EXPÉRIMENTAL . . . . .	70
3.4.1 Microscope à fond noir, ou en champ sombre . . . . .	70
3.4.2 Caméra . . . . .	71
Résolution temporelle . . . . .	71
Résolution spatiale . . . . .	72
Essais d'amélioration de notre résolution expérimentale . . . . .	73
3.4.3 Dispositif de contrôle de la température . . . . .	75
3.5 PRINCIPE D'UNE MESURE TPM . . . . .	76
3.5.1 Détection des complexes ADN/particule . . . . .	76
3.5.2 Suivi des positions . . . . .	77
Détermination de la zone d'analyse de l'objet . . . . .	78
Détermination de la position de l'objet . . . . .	78
Extraction des positions $X_i$ et $Y_i$ de l'objet . . . . .	78
3.5.3 Suivi des trajectoires . . . . .	79
Extraction du point d'ancrage de l'objet . . . . .	79

3.5.4	Correction de l'effet de dérive de la platine du microscope . . . . .	80
	Extraction de l'amplitude du mouvement l'objet . . . . .	81
3.6	FILTRES DE SÉLECTION APPLIQUÉS PAR LE LOGICIEL <i>NanoMultiplex</i> . . .	81
3.6.1	Amplitude du point d'ancrage . . . . .	82
3.6.2	Facteur de symétrie . . . . .	82
	Expression du facteur de symétrie . . . . .	82
	Valeur seuil du facteur de symétrie . . . . .	83
3.7	CORRECTIONS ET PREMIÈRES ÉTAPES DE L'ANALYSE DE DONNÉES . . . . .	85
3.7.1	Filtre sur le facteur de symétrie, $F_{Sym}$ . . . . .	87
3.7.2	Filtre sur les amplitudes, $A_{eq,i,j}$ . . . . .	87
	Exclusion des nanoparticules se collant à la surface . . . . .	87
	Exclusion des nanoparticules se décrochant de la surface . . . . .	87
3.7.3	Distribution des $A_{eq,[Moy],j}$ . . . . .	88
	Construction de la distribution des $A_{eq,[Moy],j}$ . . . . .	88
	Filtre sur les $A_{eq,[Moy],j}$ . . . . .	89
3.7.4	Valeur brute de l'amplitude du mouvement . . . . .	91
3.7.5	Correction de l'effet du moyennage vidéo . . . . .	91
	Fonction de corrélation des positions en phase $C_m(0)$ . . . . .	91
	L'erreur de pointé . . . . .	92
	Fonction de corrélation des positions hors phase $C_m(s)$ . . . . .	93
	Rapport des fonctions de corrélation . . . . .	93
	Extraction du temps de corrélation . . . . .	94
3.7.6	Correction de l'effet du moyennage vidéo . . . . .	94
	Distribution des $\mathbf{R}_{  ,i,j}$ . . . . .	94
	Distribution des $\tau_{  ,j}$ . . . . .	95
3.7.7	Filtre sur $\tau_{  }$ . . . . .	96
3.7.8	Distribution des $\mathbf{R}_{  ,i,j}$ . . . . .	97
3.7.9	Filtre sur les $\mathbf{R}_{  ,i,j}$ . . . . .	97
3.8	MESURES TPM . . . . .	100
3.8.1	Courbe de calibration . . . . .	100
3.8.2	Extraction de la contribution de l'ADN seul . . . . .	101
	Méthode minimale . . . . .	101
	Méthode de <i>Segall et al.</i> . . . . .	102
3.8.3	Extraction de la longueur de persistance . . . . .	103
	CONCLUSION . . . . .	104

## INTRODUCTION

La caractérisation de la molécule d'ADN suscite l'intérêt des chercheurs depuis de nombreuses décennies. Au vu de son rôle prédominant dans le fonctionnement de tout organisme, cet objet soulève de nombreux questionnements. Ces derniers peuvent concerner sa structure, son fonctionnement, ses propriétés physiques et chimiques, etc. Afin de répondre au mieux à ces questions, les approches expérimentales, ainsi que numériques comme nous le verrons juste après, se sont considérablement enrichies et développées. Le nombre de techniques expérimentales accessibles a explosé. La force de cette nouvelle vague de méthodes réside dans l'association de savoir-faire provenant directement du *monde de la biologie* avec une approche du *monde de la physique*.

Toute étude pertinente des **propriétés intrinsèques** à cette biomolécule, réside dans l'exploration de son comportement affranchi du maximum d'interférences possible. En effet, si nous cherchons, par exemple, à caractériser l'interaction spécifique de l'ADN avec la protéine IHF, *Integration Host Factor*, il est plus direct d'étudier l'interaction sur un système reconstitué *in vitro* dans lequel chaque partenaire protéique peut être isolé. Les recherches sur l'ADN ont donc suscité un besoin d'outils de caractérisation toujours plus précis et plus spécifiques. Ce cadre a donc favorisé le développement récent des approches à **l'échelle de la molécule unique**.

Les dernières décennies du XXème siècle, ont été le témoin de leurs premiers développements. Fin des années 70, ces méthodes furent initialement utilisées pour le suivi de molécules ou protéines dans la cellule ou sur la membrane cellulaire. En effet, le travail de Hirschfeld permet la première détection optique et acquisition d'images de protéines individuelles [Hir76]. Cette avancée est suivie de près par la détection de canaux ioniques individuels dans la membrane cellulaire [NS76]. Ces papiers sont les pionniers de la technique de suivi de molécule unique. La transposition de cette dernière sur notre objet d'intérêt qu'est l'ADN, a eu lieu en 1991 grâce aux travaux de Schafer *et al.* [Dor91]. Dans cette étude, une molécule d'ADN est fixée sur une surface de verre, attachée via une protéine RNA polymérase. L'analyse de l'étendue spatiale du mouvement brownien de la particule d'or, greffée à l'autre extrémité de l'ADN, a permis de remonter au taux d'élongation de l'ARN polymérase. En effet, la variation de l'amplitude du mouvement brownien est directement liée aux variations de la longueur apparente de la molécule d'ADN, qui elle-même est directement assujettie à l'activité de cette protéine, Fig. 3.1.

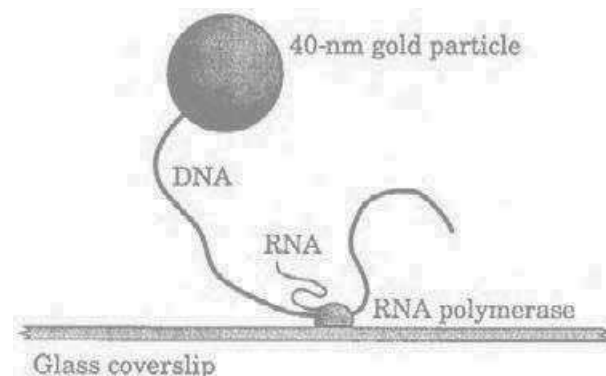


Fig. 3.1 – Extrait de [Dor91]. Schéma de la première approche de suivi du mouvement d'une molécule d'ADN attachée à une surface.

Cette technique, où le mouvement de l'ADN n'est pas contraint par l'application d'une force extérieure, est le précurseur de l'approche que nous nommons aujourd'hui *Tethered Particle Motion (TPM)*. Cette dernière est depuis largement utilisée [DZMF11]. Elle a permis d'approfondir les connaissances sur les mécanismes d'interactions ADN-protéines, la formation de boucles dans la molécule d'ADN ou encore l'enroulement, complet ou partiel, de cette dernière sur elle-même. Les mécanismes et les effets liés à la présence d'une courbure dans la molécule d'ADN commencent à être prospectés. De plus, les conditions physicochimiques jouent un rôle important et peuvent modifier la rigidité de la molécule ou encore induire une dénaturation locale. L'étude de l'ensemble de ces modifications et changements conformationnels intervenant sur la molécule d'ADN nous permet d'affiner la compréhension de certains processus biologiques. En effet, ce type d'événement intervient lors des processus de transcription et/ou de recombinaison de l'ADN.

Par la suite, les approches de **micromanipulation** se sont répandues. Ces techniques permettent d'étudier la réponse de l'ADN lorsqu'il est soumis à une contrainte. Ces contraintes peuvent être d'extension ou de torsion, notamment par pince magnétique [LAR<sup>+</sup>12]. La gamme de forces explorées par ces techniques va généralement de quelques piconewtons à plus de 100 pN [NNo8], voire à 1000 pN lors d'expériences en AFM [NNo8]. La première micromanipulation de molécules d'ADN est réalisée par Smith *et al.* en 1992 [SFB92]. La technique utilisée repose sur la combinaison de pinces magnétiques et d'un flux hydrodynamique. Ainsi, l'extension d'une molécule d'ADN a pu être mesurée en fonction de la force qui lui est appliquée. De plus, leurs résultats expérimentaux ont été comparés avec les modèles de chaînes de polymères classiques de physique statistique. Cette comparaison a permis de souligner les limites du modèle **FJC** pour décrire le comportement d'une molécule d'ADN d'environ 32  $\mu\text{m}$ . En effet, l'absence de toute corrélation d'orientation sur les positions de la chaîne ne permet pas de décrire et d'expliquer leurs résultats. Plus tard, Marko *et al.* [MS95] ont démontré que le modèle **WLC** était mieux adapté pour décrire cette réponse élastique de l'ADN sous force. En effet, ce modèle considère l'ADN comme une chaîne possédant une courbure continue, sa capacité de flexion et son entropie contribuant à son énergie.

Les approches à **l'échelle de la molécule unique** donnent accès à l'ensemble de la distribution de comportement d'une **molécule individuelle** et un accès direct à la cinétique des processus sans requérir de synchronisation. Ces mesures sont en opposition aux mesures réalisées grâce aux *approches moyenne (en : bulk)*. Ces méthodes donnent une moyenne des grandeurs caractérisées. Notre but est d'étudier des variations conformationnelles à **l'échelle de la molécule unique d'ADN**. Il faut donc utiliser une technique expérimentale adaptée et pertinente à cette échelle de mesure. Tout d'abord, nous allons décrire le **principe de la technique expérimentale** utilisée ainsi que les grandes lignes de la **préparation de nos échantillons**. Les détails du protocole sont décrits en Annexe A.1. Ensuite, nous expliciterons les routines d'**extractions des données** lors de la réalisation de nos expériences. Enfin, nous commencerons à décrire notre procédure d'**analyse des données**, commune à chacune de nos problématiques. Le script détaillé de cette analyse de données se trouve dans l'Annexe C.

### 3.1 PRINCIPE DU *Tethered Particle Motion (TPM)*

La technique de suivi de **molécule unique** utilisée se nomme en anglais *Tethered Particle Motion*, soit une traduction directe par *mouvement de la particule attachée*. La mise en place de cette technique consiste à immobiliser une des extrémités d'une molécule d'ADN sur un support solide, tandis qu'une nanoparticule, de rayon  $R_{par}$ , est attachée à l'autre extrémité, libre, de l'ADN, Fig. 3.2. Ainsi, les mouvements browniens observés au niveau de la nanoparticule nous servent de sonde afin de suivre les changements de conformation de la molécule d'ADN au cours du temps.

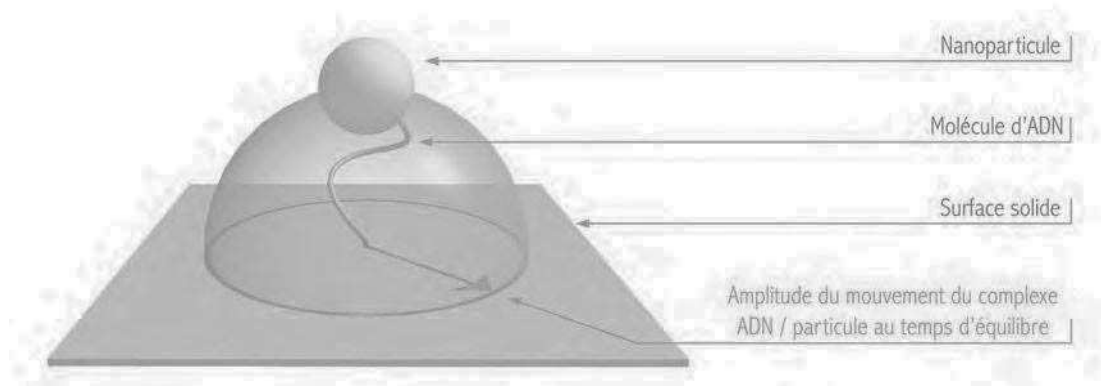


Fig. 3.2 – Schéma du principe du *Tethered Particle Motion*

Cette technique **TPM** permet de mettre en exergue la dynamique de conformations de la molécule d'ADN via la mesure de l'amplitude du mouvement du complexe ADN/particule soumis au mouvement brownien au cours du temps

L'élaboration de nos échantillons implique de suivre un protocole rigoureux pour la préparation de chacun des constituants. Les détails de ce protocole se trouvent à l'Annexe A.1, ainsi que les ajouts et modifications qui y ont été apportés durant ma thèse.

Dans la suite, seule une explication succincte du protocole est proposée (pas de détails des préparations des différentes solutions). Par contre, les tenants et aboutissants de nos différents choix liés au protocole, y sont présentés.

Les molécules d'ADN que nous cherchons à étudier sont contenues dans une *chambre* microfluidique. Celle-ci est constituée de 3 éléments principaux qui vont être assemblés, Fig. 3.3. Une fois la chambre réalisée et fermée hermétiquement, nous pouvons y introduire nos complexes ADN/particule via les ouvertures sur le dessus. Nous entamons la description des différents éléments, schématisés ci-contre, en commençant de bas en haut.



Fig. 3.3 – Schéma des composants de la chambre

### 3.1.1 Préparation du support de l'échantillon

Du fait de la taille nanométrique des objets que nous cherchons à étudier, la précision requise est très importante. De plus, comme pour toute expérience de ce type, la netteté et la fiabilité des traitements de surface sont cruciaux. En effet, il faut limiter, au maximum, la présence de contaminants qui pourrait interférer et/ou interagir avec l'ADN.

#### Préparation de la lamelle de verre

La surface solide servant de *support* à notre molécule d'ADN est une lamelle de verre de microscope. Elle est préalablement nettoyée afin d'obtenir une surface exempte de tout contaminant et présente un état de surface le plus lisse et net possible.

Une quantification de cette *propreté* peut être effectuée via une mesure dit d'angle de goutte. En effet, plus une surface est *propre*, plus elle est hydrophile. Plus une surface est hydrophile, plus une goutte d'eau va venir s'y *étaler*.

Nous pouvons voir l'évolution de l'angle de goutte au cours des étapes de nettoyage de la lamelle de verre, Fig. 3.4 ①, ② et ③. Ensuite seulement, nous **époxydons** cette surface ④.

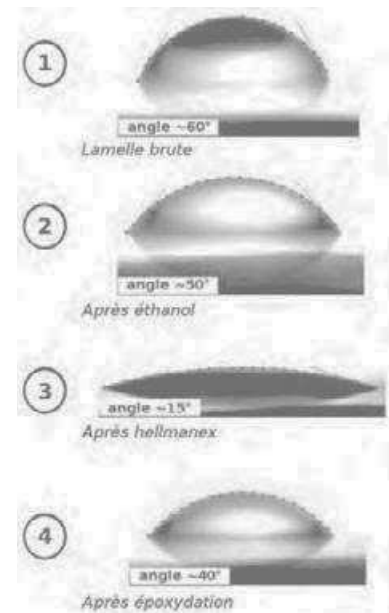


Fig. 3.4 – Mesure des angles de goutte à différentes étapes du protocole

La phase d'**époxydation** correspond à la fonctionnalisation de substrat par une molécule chimique. Cette molécule possède la fonction époxy, qui permet la liaison covalente avec les groupements amines libres des protéines que nous viendrons *tamponner* dessus. Une fois la surface époxydée, elle est mise en présence d'une solution contenant des molécules de polyéthylène glycol (PEG) ou de pluronique. Ce caractère tend à repousser stériquement les molécules d'ADN et les nanoparticules. Cela limite ainsi les adsorptions non spécifiques, et donc non désirées, sur cette dernière. Par contre, il ne faut pas que la surface soit la plus hydrophobe possible, sinon nous augmenterions la difficulté d'adsorption des protéines que nous souhaitons venir fixer sur la surface ainsi que la difficulté d'injecter de faible quantité de liquide dans notre chambre, elle-même de faibles dimensions. En effet, même à l'échelle de notre chambre micro-fluidique, les tensions de surface augmentent de façon importante.

#### Dépôt de protéines

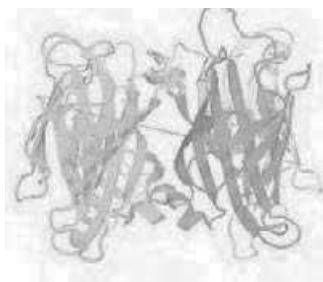


Fig. 3.5 – Neutroavidine (RCSB Protein Data Bank)

Une fois la lamelle époxydée nous pouvons venir y fixer des protéines. Dans notre étude nous travaillons avec de la **neutroavidine**<sup>a</sup> qui est constituée de 4 unités identiques, Fig. 3.5. Cette protéine possède une très forte affinité avec la biotine<sup>b</sup>. Cette interaction spécifique non-covalente met en jeu une énergie de l'ordre de  $\sim 32 \times E_{them}$  [WWPS92]. Ce couple protéine-ligand possède une des plus fortes énergies d'interaction. Les molécules de streptavidine ont une dimension  $\sim 5$  nm.

a. Neutroavidine : protéine de 64000 Da, neutre à pH 7.

b. Biotine : protéine de 244 Da.

Nous utilisons une technique dite de *micro contact printing* qui nous permet de déposer les protéines sur la lamelle époxydée. Cette étape consiste à venir apposer pendant 1 min un timbre de *Polydiméthylsiloxane*, PDMS, structuré en une succession de plots, préalablement imprégnés avec une solution contenant les protéines.

Cette mise en contact du timbre en PDMS sur la surface époxyde, transfère les protéines d'une surface à l'autre, Fig. 3.6 et 3.7. Ainsi, elles sont disposées suivant le motif défini par les plots du timbre de PDMS.

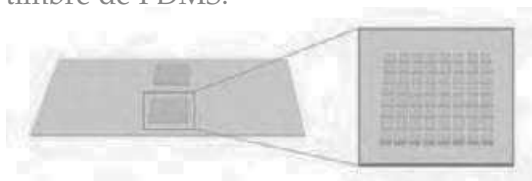


Fig. 3.6 – Schéma du dépôt de protéines

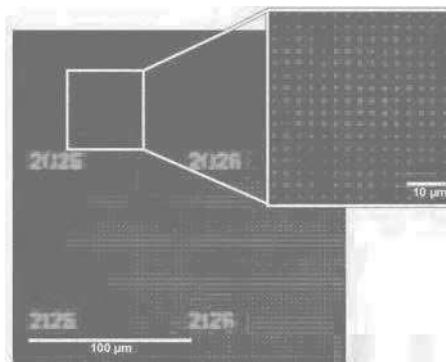


Fig. 3.7 – Visualisation par fluorescence d'un dépôt de protéines

En solution, soumis au mouvement brownien, les complexes ADN/particule que nous cherchons à ancrer sur les streptavidines, fluctuent au cours du temps. Ancré et fixé par une extrémité de l'ADN à la surface, le complexe peut explorer lors de ses mouvements une zone d'excursion limitée. Cette dernière dépend de la longueur de la molécule d'ADN ainsi que des dimensions de la nanoparticule. Un complexe ADN/particule ancré empêche, par répulsions stériques, l'accès à sa zone d'excursion à d'autres molécules.

Ainsi, la taille des zones imprimées est calibrée, selon l'estimation du rayon maximal du disque d'écrantage d'un complexe ADN/particule. La dimension de ce rayon maximal,  $R_{maximale}$ , dépend du rayon de molécule d'ADN défini suivant la **théorie Flory**,  $R_F$ , et des dimensions de la nanoparticule, selon l'expression suivante  $R_{maximale} = 2(R_{par} + R_F)$ . De plus, afin de densifier la zone de travail, tout en assurant une fluctuation libre des complexes ADN/particule, le réseau de protéines est constitué d'une succession de surface pleine de protéines de dimension  $d_{pleine} \leq R_{maximale}$ , et d'aires vides de protéines. Ces îlots de protéines sont espacés d'une  $d_{vide} \geq R_{maximale}$ , afin d'éviter tout enchevêtrement de ces complexes.

### 3.1.2 Préparation de la chambre enfermant l'échantillon

La surface servant de *couvercle* à notre chambre est une lame de verre percée de trous d'injections. Là encore, nous réalisons un traitement de surface pour la nettoyer, puis la traiter dans le but de la rendre suffisamment hydrophobe et limiter les collages non spécifiques.

#### Préparation de la lame de verre

Durant ma thèse, deux types de traitements de surface furent utilisés :

- la **thiolisation** : fonctionnalisation de la lame de verre avec une molécule chimique possédant la fonction **SH** (thiol), puis dépôt de molécule de PEG, l'Annexe A.1. Cette méthode assure un stockage des préparation, de l'ordre du mois, sans altérer la qualité de la répulsion stérique associée. Par contre,

sa réalisation nécessite deux jours de protocole. Cette méthode a été principalement utilisée lors du projet traitant de l'influence d'une courbure locale.

- l'**époxydation**, comme décrit ci-dessus pour la lamelle de verre. Cette dernière méthode ne permet pas de conserver les lames fonctionnalisées comme la méthode précédentes. En revanche, sa réalisation est plus rapide, de l'ordre de la demi-journée. Cette méthode a été utilisée lors du projet traitant de l'influence de la force ionique et de la température.

### Assemblage de la chambre

Pour finaliser la chambre, il suffit de rajouter un film de silicone d'épaisseur 0.25 mm. Ce dernier est préalablement découpé afin de laisser libre deux canaux parallèles, correspondant à un volume  $\sim 10 \mu\text{L}$  chacun. Le silicone est déposé sur la lamelle de verre. Puis la lame de verre vient refermer le tout, Fig. 3.8.

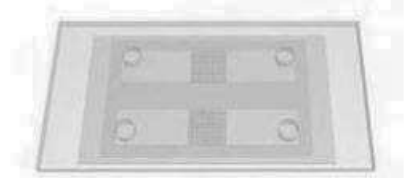


Fig. 3.8 – Schéma de la chambre microfluidique fermée

### 3.1.3 Préparation des complexes ADN/particule

Notre contenant est maintenant réalisé, il est donc essentiel de préparer les objets d'intérêt que nous cherchons à étudier.

#### ADN

Les ADNs que nous utilisons sont produits par P. Rousseau et/ou M. Salhi, via la collaboration avec l'équipe de F. Cornet du LMGM<sup>1</sup>.

Les molécules d'ADNs sont issues d'une amplification PCR (amplification en chaîne par polymérase en : *Polymerase Chain Reaction*) avec des oligomères spécifiques ce qui permet de modifier chaque extrémité de la chaîne du polymère. D'un côté, la chaîne possède une terminaison **biotinée** afin de permettre l'interaction spécifique avec les **streptavidines** disposées sur la lamelle de verre. Tandis que l'autre côté de la chaîne possède un groupement **Dig**, interagissant spécifiquement avec les anticorps **anti-Digoxygénine**<sup>2</sup> qui recouvriront la surface des nanoparticules. Les interactions de type anticorps-anti-gène sont extrêmement fortes.

Au cours de ma thèse, nous avons travaillé avec différentes tailles de molécules d'ADN. Ceci nous a permis de couvrir différents régimes de comportements des polymères. La molécule la plus longue se rapproche du régime gaussien, tandis que la plus courte,  $L \simeq 4L_p$  est plus proche du comportement d'un bâtonnet rigide.

$L_{ADN}$ bp	$L$ nm	$R_g$ nm	$D_{iff}$ si sphère dure $\mu\text{m.s}^{-1}$
2060	$\simeq 700$	$\simeq 108$	$\simeq 1.7$
1201	$\simeq 400$	$\simeq 82$	$\simeq 2.3$
575	$\simeq 200$	$\simeq 57$	$\simeq 3.3$

1. LMGM : Laboratoire de Microbiologie et Génétique Moléculaires, Toulouse-France, UMR CNRS-UPS 5100.

2. anti-Digoxygénine : anticorps de 154000 Da.



Les principales tailles d'ADN que nous avons utilisées sont répertoriées dans le tableau précédent, ainsi que l'estimation de leur rayon de giration<sup>3</sup>, Eq. 2.10. De plus, s'y trouve aussi le calcul de leur coefficient de diffusion en supposant que le rayon hydrodynamique  $R_h$  et le rayon de giration  $R_g$  sont à peu près égaux dans un milieu aqueux à 25°C, Eq. 3.1. Les séquences des ADNs utilisées sont listées en Annexe B.

### Coefficient de diffusion

Dans tout environnement, toute molécule, ADN ou nanoparticule, **diffuse**, cela est dû au **mouvement brownien**. La relation de **Stokes-Einstein** définit le coefficient de diffusion  $D$  d'une particule dans un milieu à la température  $T$  possédant un coefficient de friction  $\xi$

$$D_{diff} = \frac{k_B T}{\xi} \quad \text{Stokes-Einstein} \quad (3.1)$$

Le coefficient de friction d'une sphère de rayon  $R$  dans un milieu de viscosité  $\eta$  s'écrit  $\xi = 6\pi\eta R$ . À température ambiante, dans de l'eau  $\eta = 1.15 \times 10^{-3} \text{ Pa.s}^{-1}$ , le coefficient de diffusion d'un sphère dure de rayon  $R = 5 \mu\text{m}$ , rayon du noyau cellulaire, est égal à  $\simeq 0.04 \mu\text{m}^2.\text{s}^{-1}$ . Tandis qu'une sphère de  $R = 5 \text{ nm}$  diffusera à la vitesse de  $\simeq 40 \mu\text{m}^2.\text{s}^{-1}$ .

Les molécules du solvant diffusent. Elles entrent en collision les unes avec les autres et avec les monomères constituant le polymère. Ce phénomène entraîne des fluctuations dans les positions des objets. Tout ceci engendre aussi un changement de configuration et de conformation de l'ADN en solution. Dans le cas d'un polymère en pelote, Fig. 1.4.1,  $R$  est remplacé par le rayon hydrodynamique  $R_h$ .

### Nanoparticules

Pour permettre l'interaction spécifique de **Dig** avec l'**anti-Digoxygénine**, et donc l'accrochage de l'ADN aux nanoparticules, il est nécessaire de **fonctionnaliser** les nanoparticules avec l'anticorps. Le détail du protocole est en Annexe A.1.

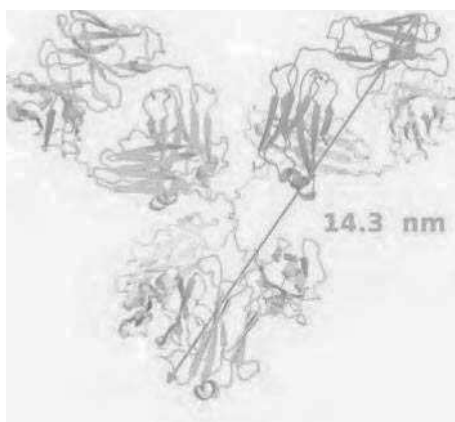


Fig. 3.9 – *Anti-Digoxygénine (RCSB Protein Data Bank)*

Les anti-Digoxygénine sont des molécules de taille plus importante que les streptavidines. En effet, la dimension de ces anticorps est d'environ 15 nm, Fig. 3.9. Au vu de leur taille, un phénomène d'intercalage entre les anti-Digoxygénine de molécule de taille inférieure est donc possible. Pour éviter ce phénomène nous ajoutons dans la solution des molécules de BSA<sup>a</sup>.

<sup>a</sup>. BSA (Bovine Serum Albumin) : molécule de 66463 Da.

3.  $R_g = \sqrt{\langle \mathbf{R}_g^2 \rangle} = \sqrt{L \times L_{p,50[\text{nm}]} / 3}$

Ces molécules permettent de combler tous les sites non occupés sur la surface des nanoparticules. Ainsi la BSA aide à bloquer les interactions non spécifiques possibles. La taille estimée des molécules de BSA est  $\sim 3$  nm.

Par défaut, les nanoparticules que nous utilisons ont un diamètre égal à  $D_{par} = 300$  nm, détail du protocole en Annexe A.1.

### Formation des complexes ADN/particule

La mise en commun d'une solution de nanoparticules fonctionnalisées par des streptavidines avec une solution d'ADNs possédant ses deux extrémités modifiées, permet la formation des complexes ADN/particule. Ce mélange est réalisé à même concentration afin de favoriser l'assemblage de type 1 ADN sur 1 particule. Il est impératif de réaliser les complexes avant de les injecter dans la chambre microfluidique. En effet, si le mélange a lieu dans la chambre elle-même il est plus difficile de contrôler le nombre de molécules d'ADN fixées par particule.

#### 3.1.4 Finalisation de l'échantillon

Avant toute chose, les deux canaux de la chambre, d'un volume  $V_{canal} \sim 10$   $\mu$ L chacun, doivent être rincés extensivement par  $\sim 100 \times V_{canal}$  de solution tampon nommée  $T_{BSA,plu}$  (détails Annexe A.1). Cette étape permet dans un premier temps, d'éliminer le surplus de protéines déposées sur la lamelle de verre. Ainsi, en moyenne, seulement une monocouche de molécule est présente. De plus, cette étape permet l'adsorption des molécules neutres de pluronique sur le verre époxydé et/ou la couche de protéine. Ces molécules adsorbées forment alors une brosse fluctuante qui produit une répulsion entropique et limite ainsi, les interactions non spécifiques d'autres objets avec la surface.

La chambre microfluidique est maintenant réalisée et prête à être utilisée. Il en est de même pour nos complexes ADN/particule en solution. Nous pouvons injecter la solution de complexes ADN/particule dans chacun des canaux. Ensuite, le tout est laissé à incuber toute une nuit. Ce temps d'incubation permet la diffusion des complexes jusqu'aux sites de fixations que constituent les îlots de streptavidines sur la surface.

### Diffusion de nos complexes ADN/particule dans la chambre microfluidique

L'estimation du temps de diffusion minimal de nos complexes pour parcourir la distance définie par l'épaisseur du silicone  $e_{silicon} = 0.25$  mm, est régie par le coefficient de diffusion des particules de diamètres  $D_{par} = 300$  nm. Le calcul du temps d'incubation minimal est le suivant :

$$\tau_{minimal} = \frac{e_{silicon}^2}{6.D_{iff}} \sim 3h \quad (3.2)$$

## 3.2 TPM-ON-A-CHIP

Le principe du **TPM-on-a-chip** est d'obtenir un haut débit d'informations sur nos objets ADN/particule, en seulement une expérience. L'idée est donc de **paralléliser** les mesures de **suivi de molécule unique**, Fig. 3.10. Pour cela,

notre protocole expérimental permet la **multiplication** des *objets-TPM* : complexes ADN/particule accrochés par une extrémité à la surface solide via l'interaction streptavidine-biotine. En effet, le motif de protéine que nous avons déposé via le tampon en PDMS permet la fixation d'un, et un seul, complexe ADN/particule par îlot. Notre surface de travail est donc remplie d'*objets-TPM*.

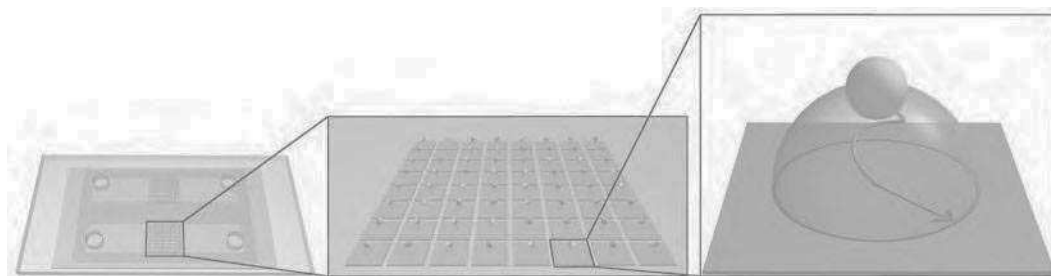


Fig. 3.10 – Schéma du TPM-on-a-chip

Notre méthode expérimentale nous permet d'extraire la **projection en 2 Dimensions** du mouvement de la particule. En effet, nous observons notre échantillon selon le **plan (XY)**, nous ne collectons pas d'informations quantitatives sur la position selon l'axe Z. Les interactions fortes et spécifiques du couple **Dig-anti-Digoxygénine** permettent de former des complexes ADN/particule. Grâce aux interactions non-covalentes protéine-ligant, l'autre extrémité biotinée de l'ADN vient s'**ancrer** sur les streptavidines fixées sur la surface époxydée. En solution, les complexes ADN/particule sont soumis au **mouvement brownien**.

Le mouvement brownien de chaque complexe ADN/particule est contenu dans un **volume d'excursion**. Les dimensions de ce volume dépendent de la longueur de l'ADN ainsi que du rayon de la nanoparticule. De plus, chaque objet est confiné dans une **demi-sphère** due à la présence de la surface.

### 3.3 POINTS FORTS DU TPM-ON-A-CHIP

L'avantage principal de la technique TPM est que cette approche n'influe et ne réduit pas, ou très peu, les degrés de liberté de conformation de la chaîne d'ADN. En effet, aucune contrainte ou force extérieure significative n'est exercée sur la biomolécule. Ainsi, cette méthode donne accès aux conformations de la chaîne.

La présence de la surface époxydée, contraint le complexe ADN/particule à explorer seulement une demi-sphère. La présence de cette surface induit une force de répulsion, excluant ainsi l'espace correspondant à l'autre demi-espace se trouvant sous la lamelle. Autrement dit, l'ensemble ADN/particule tend à être étiré vers le haut. La **force effective** résultant de ce phénomène de **volume exclu**, correspond à une force d'étirement verticale de la particule sur l'ADN. L'estimation de cette force est donnée dans l'article de Segall *et al.* [SNP06].

La mesure de la force effective dépend donc de la longueur de l'ADN ainsi que du rayon de la nanoparticule. Si nous supposons être en régime gaussien, l'expression de la force effective s'écrit telle que :

$$\langle \mathbf{F}_{\text{eff}} \rangle = \frac{k_B T}{\sqrt{\pi(LL_p/3)}} \cdot \frac{1 - \exp[-N_R^2]}{\text{erf}(N_R)} \quad (3.3)$$

Où  $\text{erf}(x)$  est la fonction d'erreur. Le terme  $N_R$  correspond au **nombre d'excursions** de l'ADN. Ce nombre permet de définir si la contribution de l'ADN ou celle de la particule, domine le mouvement brownien de complexe ADN/particule. Si  $N_R < 1$ , le mouvement de l'ADN prédomine sur celui de la particule. Tandis que si  $N_R > 1$ , la contribution de la particule au mouvement brownien est la plus importante. L'expression de **nombre d'excursions** est la suivante :

$$N_R = \frac{R_{\text{par}}}{\sqrt{(LL_p/3)}} \quad (3.4)$$

Considérons un ADN de  $L_{\text{ADN}} = 2060$  bp soit  $L \sim 700$  nm, de longueur de persistance  $L_p = 50$  nm, fixé à une particule  $R_{\text{par}} = 150$  nm. Le **nombre d'excursions** est égal à 1.4. L'estimation de la **force effective** exercée par la particule sur l'ADN est égale à  $\sim 2 \times 10^{-2}$  pN soit  $\sim 20$  fN. La valeur force exercée sur l'ADN lors de nos expériences TPM est quasi nulle,  $F_{\text{eff}}$  est inférieure au piconewton. Pour étudier le mouvement de l'ADN, il sera par contre nécessaire de soustraire la contribution de la particule au mouvement de l'ensemble ADN/particule. Nous verrons le détail de ce calcul lors de l'analyse des données TPM, partie 3.8.2

De plus, les expériences TPM disposent d'une mise en œuvre facile. Le matériel reste peu coûteux et la mise en place est assez rapide, en comparaison avec d'autres techniques expérimentales.

### 3.4 DISPOSITIF EXPÉRIMENTAL

La précision des mesures est intrinsèquement liée à la technique et au dispositif de mesure. Le développement et la précision des caractéristiques du dispositif expérimental vont permettre de mettre en lumière les points forts de notre approche ainsi que ses limites techniques. Ceci va nous permettre de définir les résolutions spatiales et temporelles de notre TPM-on-a-chip.

#### 3.4.1 Microscope à fond noir, ou en champ sombre

Afin d'optimiser la qualité du signal, nous exploitons la technique de microscopie en champ sombre ou dit à fond noir. Notons aussi que nous utilisons un microscope inversé, l'objectif se trouve sous l'échantillon, Fig. 3.11 ①, ②. En effet, cette configuration rend plus aisée les injections en cours de manipulation grâce à la présence de trous d'injections placés sur le dessus de la chambre microfluidique.

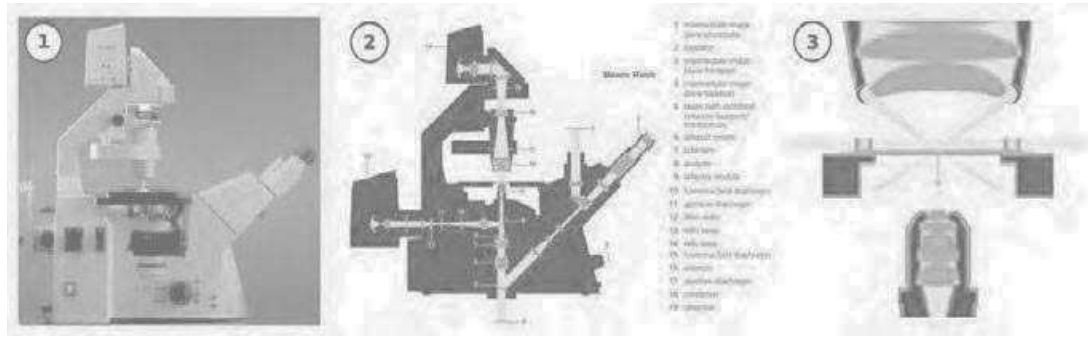


Fig. 3.11 – ① Photo du Microscope Zeiss Axiovert 200 utilisé. ② Schéma du trajet du faisceau lumineux dans le microscope, extrait des documents Zeiss. ③ Zoom du trajet du faisceau lumineux en incidence rasante lors de la traversée de l'échantillon

La microscopie à fond noir permet d'étudier des objets dont les dimensions ou la transparence les rendent difficilement discernables habituellement. Cette technique de microscopie consiste, non pas à illuminer notre échantillon en lumière directe comme sur un microscope classique, mais à éclairer l'échantillon par un faisceaux lumineux fortement incliné, Fig. 3.11 ③. L'utilisation d'un obturateur placé au niveau du diaphragme d'un condenseur, permet d'éclairer l'échantillon avec un cône de lumière doté d'un fort angle. L'objectif, muni d'une ouverture numérique inférieure à celle du condensateur, collectera uniquement la lumière diffusée par les objets de l'échantillon.

Cette technique permet d'améliorer le contraste de l'échantillon. L'intensité du signal est fortement améliorée par rapport à la microscopie en fond clair. De plus, ceci permet de nous affranchir d'un signal discontinu lié à l'utilisation des quantum-dots<sup>4</sup>, ou d'une baisse de l'intensité des particules fluorescentes induite par le photoblanchiment<sup>5</sup> des fluorophores encapsulés.

Nous utilisons un condensateur ( $ON_{cond} = 0.55$ ), et un objectif de microscope  $G_{obj} = \times 32$  ( $ON = 0.4$ ) auquel est ajouté un tube d'agrandissement  $G_{mag} = \times 1.6$ .

### 3.4.2 Caméra

#### Résolution temporelle

La caméra utilisée est une CMOS camera Dalsa Falcon 1.4M100 qui nous permet une fréquence d'acquisition de  $f_{acq} = 25$  Hz. Ce taux d'enregistrement nous permet l'acquisition d'une image toutes les  $\tau_{acq} = 40$  ms. De plus, le temps d'exposition d'une image est de  $\tau_{ex} = 40$  ms. Enfin, toute mesure nécessite l'accumulation d'un nombre suffisant de mesures de positions décorréliées de notre système. Car nous cherchons à mesurer des amplitudes du mouvement qui sont elles-même des moyennes temporelles, Eq. 3.10. Autrement dit, nous devons accumuler un nombre suffisant d'images sur la durée du moyennage  $T_{av}$ . Ceci diminue donc notre résolution temporelle.

4. Quantum-dot (boîte quantique) : objet nanométrique tridimensionnel semi-conducteur. Soumis à une énergie d'excitation (faisceaux lumineux) cet objet à la capacité d'émettre un photon pour se désexciter. L'émission de photon dépend du temps caractéristique d'excitation et de désexcitation. Un quantum dot émet donc une lumière discontinue.

5. Photoblanchiment : caractérise la perte de la capacité de fluorescence suite à une exposition prolongée et/ou à une forte lumière.

Comme détaillé dans l'article de 2010 [MTB<sup>+</sup>10], le choix de la valeur de  $T_{av}$  est crucial et dépend des paramètres de relaxation du système étudié. En effet, d'une part si  $T_{av}$  est trop faible, ceci induit une augmentation du bruit, erreur statistique, sur la mesure, ce qui implique une mauvaise détection et précision des changements d'amplitudes du mouvement. Dans le cas contraire, lorsque la valeur de  $T_{av}$  est trop important, ceci diminue la précision car les positions sont alors trop moyennées. Afin de respecter ces contraintes nous avons donc choisi  $T_{av} = 2$  s.

Ainsi, nous pouvons nous intéresser et distinguer la dynamique d'un événement possédant une longue durée de vie. Autrement dit une durée de vie supérieure à notre temps d'acquisition multiplié par le nombre d'images nécessaire afin d'obtenir des mesures de position décorréliées de notre système, par exemple, l'accrochage et la fixation d'une protéine sur l'ADN [DSC<sup>+</sup>14].

Intéressons-nous à un événement ayant une durée de réalisation inférieure à 40 ms, par exemple la fermeture d'une bulle de dénaturation. Selon les prédictions numériques, la fermeture complète d'une bulle de 40 à 100 bp s'effectue sur seulement 0.1 à 100  $\mu$ s [DDPM13].

Si nous nous plaçons dans des conditions où *la bulle se ferme*, nous ne pourrions pas voir la progression de la fermeture de la bulle via notre acquisition. En effet, une image sera la moyenne des conformations que l'ADN aura adopté durant les 40 ms correspondant au temps d'exposition. Cette durée est supérieure au temps de l'événement considéré. Nous ne pourrions pas acquérir la progression de l'événement image par image, nous ne verrons donc pas entièrement la dynamique de cette événement avec notre système.

Notre résolution temporelle est de l'ordre de grandeur de  $T_{av}$ . Ainsi, par le fait de moyenniser des positions sur la durée  $T_{av}$ , seuls des durées d'événement se réalisant sur un temps plus grand que  $T_{av}$  peuvent être détectées avec précision. Pour l'observation d'événement à durée de vie courte, donc inférieure à  $T_{av}$ , il est nécessaire de diminuer la dynamique. Pour ce type d'événement des procédés de ralentissement de la dynamique du système biologique seront à envisager dans l'avenir.

### Résolution spatiale

La caméra utilisée possède 1.4 millions de pixel, leur taille est de  $S_{pixel} = 7.4 \times 7.4 \mu\text{m}$ . Le rapport d'agrandissement entre l'objet réel et ce qui est projeté sur le capteur de la caméra est de  $G_{cam} = 48$ .

La caméra utilisée est monochrome, ainsi tout objet est enregistré et codé en niveau de gris. Le codage maximal des niveaux de gris est de 255 à 0, la valeur 0 est associée au noir.

Le grossissement de notre dispositif nous permet de calculer la calibration de notre système optique, correspondant à la taille réelle d'un objet de l'échantillon qui est visualisé sur un pixel de la caméra. La taille d'un objet correspondant à un pixel est de :

$$S_{objet} = \frac{S_{pixel}}{M_{cam}} = 154 \text{ nm} \quad (3.5)$$

Or, il est nécessaire de prendre en compte le **phénomène de diffraction**, dû au caractère ondulatoire de la lumière. La lumière émise par un points source ponctuel puis passant à travers une petite ouverture optique, est diffractée sur le plan objet en une succession de taches étendues et concentriques nommées anneaux d'Airy, Fig. 3.12.

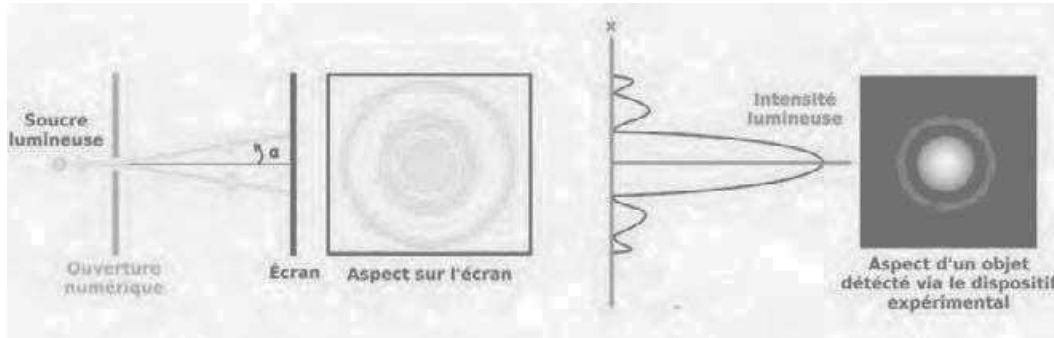


Fig. 3.12 – Illustration du phénomène de diffraction lors du passage d'un faisceau lumineux à travers une ouverture de petite dimension, de l'ordre de grandeur de la longueur d'onde

Le rayon minimal des anneaux d'Airy est défini tel que :

$$R_{\text{objet}} = \frac{0.61 \cdot \lambda}{ON} \quad \text{avec} \quad ON = n \cdot \sin(\alpha) \quad (3.6)$$

Soit  $\lambda$  la longueur d'onde de la lumière avec laquelle l'échantillon est observé ;  $ON$  l'ouverture numérique du dispositif optique avec lequel la lumière est collectée ;  $n$  l'indice de réfraction du milieu et  $\alpha$  le demi-angle d'ouverture de l'objectif.

Ce rayon définit donc la résolution latérale, imposée par les lois de l'optique, de notre système de microscopie appelée limite de la diffraction ou loi d'Abbé. Une source lumineuse ponctuelle du spectre visible,  $\lambda \in [400 : 800]$  nm passant à travers une ouverture numérique de 0.4, sera donc observée tel un disque de rayon  $R_{\text{objet}} \sim 1220$  nm. Ce point source lumineux est donc, à minima, défini par une tache d'un diamètre de 16 pixels. De plus, cette valeur est à convoluer par la taille du rayon de la nanoparticule utilisée dans l'expérience.

L'ensemble de notre dispositif expérimental nous permet d'obtenir un champ d'observation de dimensions  $215 \mu\text{m} \times 160 \mu\text{m}$  lors d'une seule acquisition.

### Essais d'amélioration de notre résolution expérimentale

Pour améliorer la précision sur notre mesure de la dynamique du système, nous avons cherché à diminuer la taille des nanoparticules utilisées. En effet, la taille des nanoparticules de 150 nm de rayon est plus importante que le rayon de giration de nos ADNs. La dynamique est donc dominée par la dynamique des nanoparticules qui possèdent le coefficient de diffusion le plus lent. Les processus et/ou événements intervenants sur l'ADN avec une vitesse plus rapide que l'ordre de vitesse de la nanoparticule seront ralentis. Nous avons donc prospecté des tailles de nanoparticules, répertoriées dans le tableau suivant, afin d'éviter ce phénomène.

Matériaux	$D_{par}$ $\mu\text{m}$	$R_{par}$ $\mu\text{m}$	$D_{iff}$ si sphère dure $\mu\text{m.s}^{-1}$
Latex	300	150	$\sim 1.3$
Latex	200	100	$\sim 2.0$
Or	80	40	$\sim 4.8$
Or	40	20	$\sim 9.5$

Tout d’abord, sur la même base du protocole de fonctionnalisation et restant avec des nanoparticule de latex, nous avons testé des objets de diamètre inférieur, passant de  $D_{par} = 300$  nm à 200 nm. Cette diminution de la taille des nanoparticules, diminue la section efficace de diffusion de l’objet. Cette section efficace définit l’intensité lumineuse transmise par l’objet. Le gain en dynamique via l’utilisation des nanoparticules de  $D_{par} = 200$  est trop faible par rapport à la diminution du signal.

Nous avons donc prospecté des nanoparticule donc les sections efficaces sont plus importantes que celles des nanoparticules de latex  $D_{par} = 300$ . Pour cela nous nous sommes tournés vers des nanoparticules d’or. Ces dernières possèdent deux avantages principaux. Tout d’abord, leur synthèse est très bien contrôlée ce qui permet d’obtenir des objets très homogènes en taille, l’erreur relative sur le diamètre  $< 8\%$ . De plus, les nanoparticules d’or possèdent une résonance plasmon dans le spectre visible. En effet, lorsqu’une particule métallique est soumise à un champ électromagnétique dont la longueur d’onde est largement supérieure à la taille de cette dernière, tous les électrons libres de la bande de conduction de l’objet subissent le même champ électromagnétique. Ainsi, les électrons oscillent collectivement et en phase, Fig. 3.13.

Lorsque la fréquence de l’onde incidente correspond à la fréquence propre de ces oscillations, il se produit un phénomène de résonance, appelé résonance de plasmon de surface. Cette résonance a lieu dans le domaine du visible, seulement pour l’or, le cuivre et l’argent

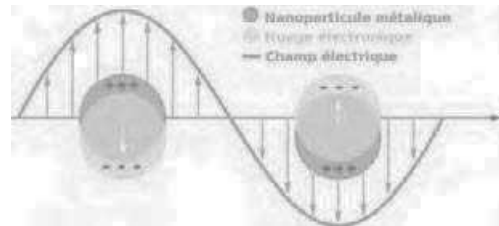


Fig. 3.13 – Illustration de la résonance plasmon

Comme nous pouvons le voir sur la Fig. 3.14, seules les nanoparticules d’or  $D_{par} = 80$  possèdent une section efficace de diffusion intéressante pour nos mesures. L’utilisation de ces dernières avec les mêmes caractéristiques du dispositif expérimental, induit un meilleur rapport signal sur bruit. Or, dû à la diminution de la taille des nanoparticules, le temps de relaxation des complexes ADN/particule est fortement réduit. Il est donc nécessaire d’améliorer la fréquence d’acquisition de la caméra. En effet, si nous restons avec les paramètres actuels, nous perdons en précision spatiale due à l’augmentation du nombre de positions moyennées lors de la réalisation d’une image. Nos essais avec une nouvelle caméra ont dû être laissés de côté suite à un problème important de communication entre cette dernière et notre logiciel d’analyse *Nanomultiplex*.



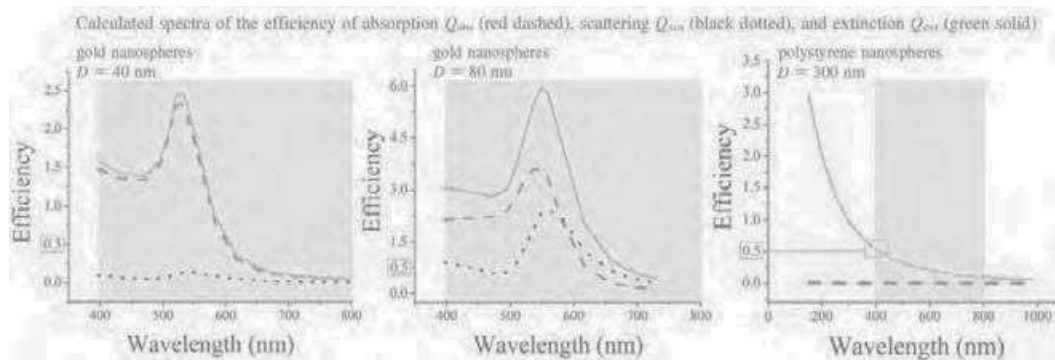


Fig. 3.14 – Spectres définissant l'efficacité des sections efficaces de diffusion, courbe noire, d'absorption, courbe rouge, et d'extinction, courbe verte, pour des nanoparticules d'or  $D_{par} = 80, 40$  et de latex  $D_{par} = 200$ , adapté de [JLESESo6]

Afin de résoudre ce problème, un investissement assez important en temps est nécessaire. Nous avons dû stopper ces essais en cours de développement, afin de continuer nos avancées sur ce sujet de thèse. Néanmoins, l'utilisation de ces nanoparticules reste l'option principale afin de nous permettre de diminuer le temps de relaxation du système.

Les différents protocoles, en fonction du matériel utilisé ainsi que de leur taille, que nous avons réalisés sont disponibles en Annexe A.1.

### 3.4.3 Dispositif de contrôle de la température

Les propriétés physico-chimiques de la molécule d'ADN, et notamment sa flexibilité de courbure sont fortement dépendantes de la température du milieu. En effet, une variation de la température influe sur la longueur de persistance globale et donc sur la mesure de la distance bout-à-bout de la molécule d'ADN [GH68, GKV11]. Il est donc primordial de contrôler précisément la température. En effet, nous cherchons à réaliser des mesures fines de changement de conformation de l'ADN induites par des contraintes qui sont indépendantes de  $T$ .

Pour cela nous utilisons un Peltier *Physitemp* capable de tenir la gamme de températures suivantes  $T_{Peltier} \in [-10 : 100]^\circ\text{C}$ .

Afin d'isoler le Peltier de la platine du microscope nous avons prospecté la capacité de conduction thermique de certains matériaux.

Matériau	$\sigma_{Thermique}$ à $20^\circ\text{C}$ en $\text{W.m}^{-1}.\text{K}^{-1}$
Teflon	0.25
PDMS	0.15
Contre plaqué	0.11
Carton	0.07
Liège	0.04
Polyester expansé	0.036

TABLE 3.1 – Coefficient d'échange thermique de quelques matériaux

Le carton est l'un des matériaux possédant l'un des plus faibles coefficients d'échange thermique, Table. 3.1. De plus, ce matériau est fortement répandu, peu coûteux et surtout il est extrêmement facile de confectionner une pièce aux formes et dimensions souhaitées.

Afin d'isoler le Peltier de la platine du microscope et de le fixer convenablement, nous avons fabriqué une succession de pochoirs cartonnés qui s'emboîtent les uns sur les autres, comme visible sur les Fig. 3.15 et 3.16. De plus, nous ajoutons un couvercle transparent afin de réduire les variations de températures induites par flux convectif.



Fig. 3.15 – Photographie du Peltier Physitemp



Fig. 3.16 – Schéma du dispositif de contrôle de la température

### 3.5 PRINCIPE D'UNE MESURE TPM

Suite à l'incubation pendant une nuit de nos complexes ADN/particule, nous rinçons extensivement notre échantillon par un volume  $\sim 100 \times V_{canal}$ . Ceci permet d'éliminer tout objet non accroché présent en solution. Nous pouvons donc faire la mise au point sur nos objets d'intérêt.

#### 3.5.1 Détection des complexes ADN/particule

Rappelons que ce sont les nanoparticules qui nous servent de sonde pour suivre le mouvement brownien de l'ADN. Ce sont donc ces dernières que nous distinguons sur la Fig. 3.17 comme objets diffusants. Le système optique du microscope diffracte le signal lumineux diffusé par les points sources que constituent les nanoparticules, ceci les fait apparaître telles des taches et non plus seulement des points.

Communément, le champ d'observation lors de nos expériences de TPM-on-a-chip, se présente sous la forme suivante :

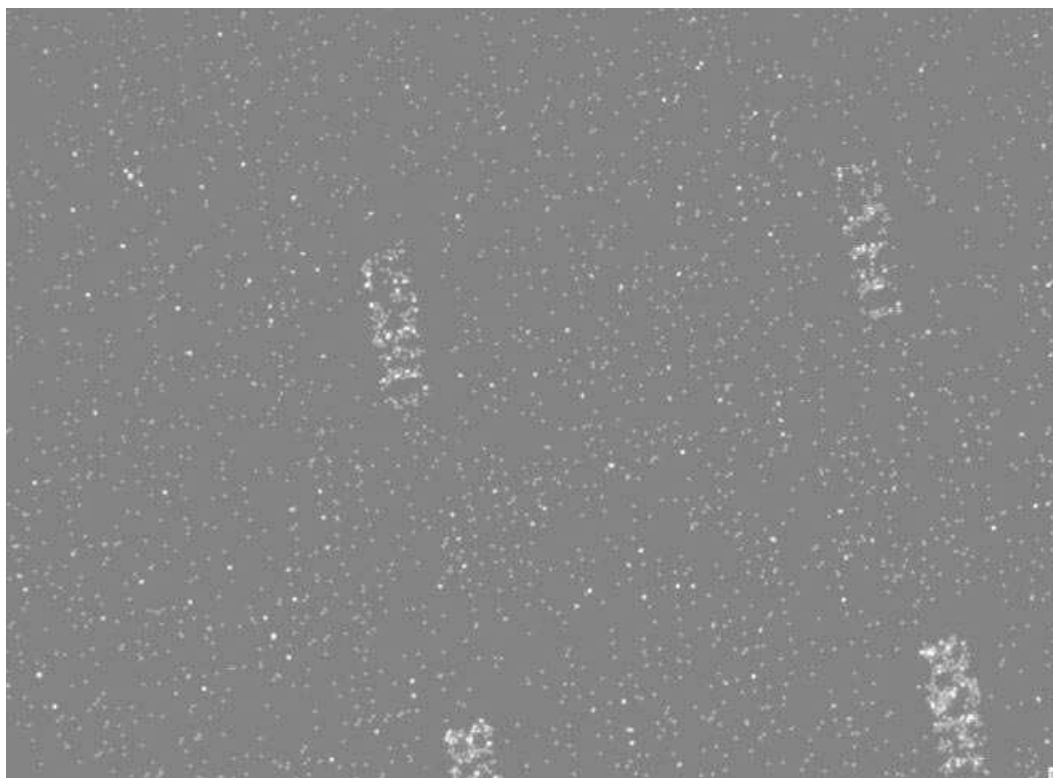


Fig. 3.17 – Champ d'observation en champ sombre, soit  $215 \times 160 \mu\text{m}$ . Les complexes ADN/particule sont réalisés avec  $L_{\text{DNA}} = 575 \text{ bp}$ ,  $R_{\text{par}} = 150 \text{ nm}$ , dans des conditions salines proches des conditions physiologiques. Chaque point blanc correspond à une nanoparticule. Les taches de taille plus importante, non homogène et/ou asymétrique correspondent à des agrégats de quelques particules. Les zones de forte densité de nanoparticules, où l'on peut distinguer des numéros, correspondent à des balises pour se repérer sur le motif.

L'arrangement spatial des complexes ADN/particule vient de l'arrangement régulier du motif de protéines transposé via le timbre en PDMS. De plus, la quantité de complexes ADN/particule individuellement fixée sur la surface est impressionnante. Sur cette image le nombre brut des objets présents avoisine les 1500, et ce pour une seule expérience. Mais, tous ces complexes ne sont pas valides, comme nous allons le voir.

### 3.5.2 Suivi des positions

Les positions et le mouvement des nanoparticules sont alors étudiés à l'aide d'un logiciel intégrant un certain nombre de calculs, (via le logiciel *Nanomultiplex* commandé à la compagnie privée *Magellium*). Ce dernier va nous permettre, en plusieurs étapes, de déterminer des paramètres dynamiques des nanoparticules : leur position au cours du temps, leur amplitude de mouvement, leur point d'ancrage, le facteur d'asymétrie de leur mouvement.

En premier lieu il est nécessaire de définir, déterminer puis sélectionner les objets d'intérêt sur le champ d'observation.

### Détermination de la zone d'analyse de l'objet

Afin de déterminer les zones d'intérêt, l'image est seuillée, (*commande : Threshold*), afin d'accentuer le contraste entre les objets diffusants et le fond noir. En fonction de la valeur du seuil, un plus ou moins grand nombre de niveaux de gris sera mis à une valeur numérique de 0, afin de ne pas être pris en compte, Fig. 3.18.

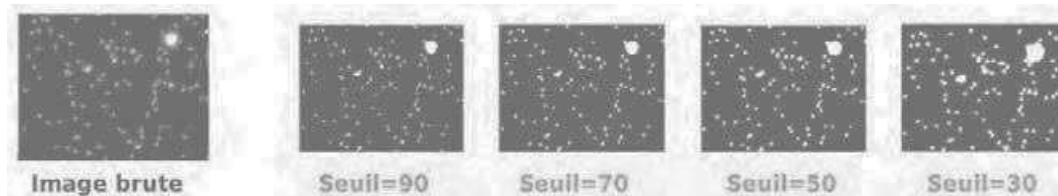


Fig. 3.18 – Visualisation de l'effet du seuil, proche d'un agrégat pour souligner le phénomène. Image brute est réalisée avec l'intensité lumineuse de la lampe du microscope égale à 8 V, nous y distinguons l'hétérogénéité du fond noir. Images seuillées avec différentes valeurs.

La valeur du seuil dépend de la qualité du fond, du fond noir et de l'intensité lumineuse incidente. Pour une valeur de seuil trop grande, un grand nombre de niveau de gris sont mis à zéro, ceci induit une altération des objets observés. Les objets sont tronqués, dû au peu de nuances de gris conservées. Une valeur de seuil trop faible, implique là encore une altération des objets. Ces derniers sont déformés, nous observons parfois la fusion de deux objets spatialement proches.

Une fois le seuil convenablement choisi il est conservé pour toute la durée de l'expérience sur le même échantillon afin de garder le même signal sur bruit.

Cette étape permet de distinguer les objets du fond. Autour de chaque objet une zone d'intérêt réduite est créée, (*en : R.O.I, Region Of Interest*). Les tailles de ces zones sont définies par des boîtes de taille variant, généralement, de 10 à 20 pixels de côté et centrées sur les particules (*commande : Bounding Box Size (pixels)*).

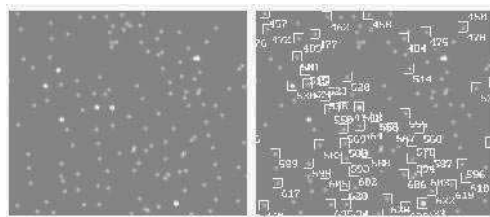


Fig. 3.19 – Visualisation des R.O.I.

Ainsi, Fig. 3.19, chaque complexe ADN/particule est repéré afin d'être traité et analysé **individuellement**.

De plus, cette étape permet de ne pas traiter l'ensemble de l'image mais seulement des sous-parties, ce qui augmente grandement la vitesse d'analyse.

### Détermination de la position de l'objet

Sur chaque image, le centre de masse de l'objet est mesuré. Ce dernier est défini comme le barycentre pondéré de l'intensité de chaque pixel constituant le signal lumineux de type gaussien qui correspond à nos nanoparticules. Les coordonnées du centre de masse correspondent aux **positions absolues**  $X_i$  et  $Y_i$  de la particule à chaque image  $i$ .

### Extraction des positions $X_i$ et $Y_i$ de l'objet

Pour chacun des complexes ADN/particule, le logiciel calcule les positions  $X_i$  et  $Y_i$  au cours du temps, soit pour chaque image toutes les 40 ms. Ces positions sont collectées dans des fichiers individuels correspondant à chacun de nos objets.

### 3.5.3 Suivi des trajectoires

La grandeur qui nous intéresse est la **distance bout-à-bout de notre molécule d'ADN**. Pour cela nous devons remonter à la trajectoire de la particule au cours du temps par rapport à un point de référence. Ce point de référence est bien entendu le **point d'ancrage de notre ADN sur la surface**.

#### Extraction du point d'ancrage de l'objet

Tel un chien attaché à un piquet par une laisse, l'ensemble de l'espace explorable et accessible au chien est centré sur le point fixe qu'est le piquet. Pour déterminer le point d'ancrage il est donc nécessaire de connaître les limites du **volume d'excursion** de notre système. Autrement dit, il est nécessaire d'avoir laissé assez de temps au système pour explorer l'espace qui lui est accessible. Ce qui signifie que le point d'ancrage de nos complexes ADN/particule ne peut être déterminé sur une seule image, car une image correspond à un instantané du mouvement.

Le calcul du point d'ancrage se fait donc sur une durée largement supérieure au temps de corrélation des positions. En pratique, le calcul du point d'ancrage est effectué sur toute la durée de notre acquisition grâce à une fenêtre *glissante* de  $Nbr_{image}$  images.

Du fait de notre dispositif expérimental une image est prise toutes les  $\tau_{acq} = 40$  ms, ce qui correspond à un temps suffisant afin que les positions soient décorrélées. Le temps suffisant est estimé à  $\tau_{Anc} = 2$  secondes d'acquisition, (*commande : Sliding Window Size (s)*), à une fréquence d'acquisition de 25 images par seconde [PDT<sup>+</sup>04]. Les temps de relaxation de nos complexes,  $\tau_{||}$ , varient de  $\sim 20$  à  $\sim 60$  ms en fonction de la longueur de la molécule d'ADN utilisée [PDT<sup>+</sup>04]. Le calcul du point d'ancrage est réalisé sur une durée allant de  $\sim 50$  à 100 fois  $\tau_{||}$ .

$$\begin{aligned} Nbr_{images} &= \frac{\tau_{Anc}}{\tau_{acq}} = \frac{2 \times 10^3 \text{ ms}}{40 \text{ ms}} \\ Nbr_{images} &= 50 \text{ images} \end{aligned} \quad (3.7)$$

Afin de respecter cet intervalle nous prenons  $Nbr_{images} \Rightarrow 50 + 1$  images.

Les coordonnées du **point d'ancrage**  $X_{anc,i}$  et  $Y_{anc,i}$ , sur une fenêtre glissante de  $Nbr_{images} = 51$ , s'expriment :

$$\begin{aligned} X_{anc,i} &= \frac{1}{51} \sum_{k=i-25}^{i+25} X_k \\ Y_{anc,i} &= \frac{1}{51} \sum_{k=i-25}^{i+25} Y_k \end{aligned} \quad (3.8)$$

Le calcul du point d'ancrage est directement réalisé dans le logiciel *NanoMultiplex*, les positions de ce dernier sont archivées dans des fichiers individuels correspondant à chacun de nos objets.

### 3.5.4 Correction de l'effet de dérive de la platine du microscope

Il existe une dérive de la platine du microscope au cours du temps Fig. 3.20. Une méthode pour corriger ce *drift* est de corriger les coordonnées, les **positions absolues**  $X_i$  et  $Y_i$  pour chaque image  $i$  en soustrayant les coordonnées du **point d'ancrage**  $X_{anc,i}$  et  $Y_{anc,i}$ . Ainsi, nous obtenons les **positions relatives**,  $X_{cor,i}$  et  $Y_{cor,i}$ , corrigées de la dérive de la platine du microscope.

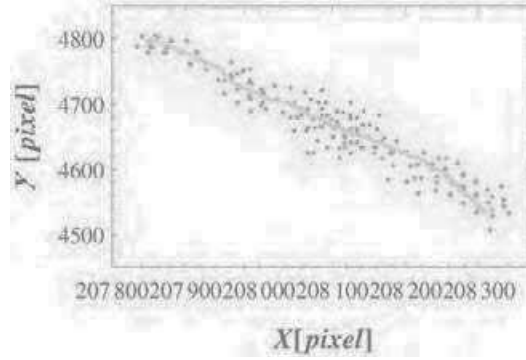


Fig. 3.20 – Illustration de la dérive de la platine du microscope sur un objet fixe. En vert les **positions absolues** des complexes ADN/particule, en orange les positions du **point d'ancrage**

Le calcul se fait sur une fenêtre glissante de  $Nbr_{images} = 51$ . L'expression des **positions relatives** au point d'ancrage,  $X_{cor,i}$  et  $Y_{cor,i}$ , s'expriment :

$$\begin{aligned} X_{cor,i} &= X_i - X_{anc,i} \\ Y_{cor,i} &= Y_i - Y_{anc,i} \end{aligned} \quad (3.9)$$

La représentation Fig. 3.21 ① ci-dessous, présente les positions  $X_i$  et  $Y_i$  au cours d'une minute d'acquisition ainsi que les positions du point d'ancrage  $X_{anc,i}$  et  $Y_{anc,i}$ . L'étalement des positions du point d'ancrage montre l'effet de dérive. Les positions  $X_{cor,i}$  et  $Y_{cor,i}$  en Fig. 3.21 ②, correspondent aux positions après application des corrections, elles sont donc centrées en (0,0).

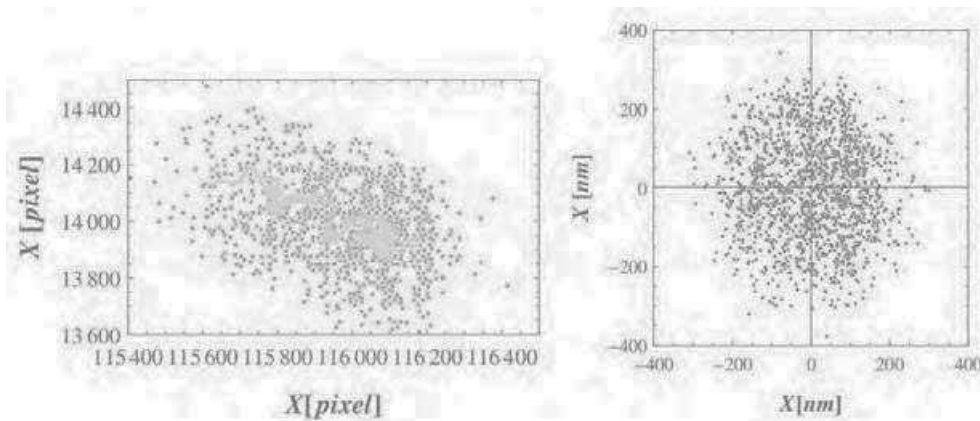


Fig. 3.21 – ① En vert les **positions absolues** des complexes ADN/particule, en orange les positions du **point d'ancrage**. ② Les positions **positions relatives**. (ADN  $L_{ADN} = 1201$  bp,  $R_{par} = 150$  nm)

Cette opération est directement réalisée dans le logiciel *NanoMultiplex*. Comme nous allons le voir, ce calcul intervient lors de la détermination de l'amplitude du mouvement au cours du temps,  $A_{eq,i}$ , des complexes ADN/particule.

### Extraction de l'amplitude du mouvement l'objet

Nous cherchons à connaître l'**amplitude du mouvement en  $i$** , soit  $A_{eq,i}$ , Fig. 3.22 ①. Autrement dit la valeur de la moyenne quadratique des distances entre le point d'ancrage et la position de la nanoparticule  $X_{cor,i}$  et  $Y_{cor,i}$  à chaque temps  $i$ . Là encore, le calcul se fait sur une fenêtre glissante de 51 images, et s'exprime tel que :

$$A_{eq,i} = \left[ \frac{1}{n} \sum_{k=i-n}^{i+n} (X_i^2 + Y_i^2) \right]^{1/2} \quad (3.10)$$

Le calcul est réalisé dans le logiciel, les amplitudes du mouvement des complexes ADN/particule sont collectées dans les fichiers individuels correspondant à chacun de nos objets, à la suite des autres données.

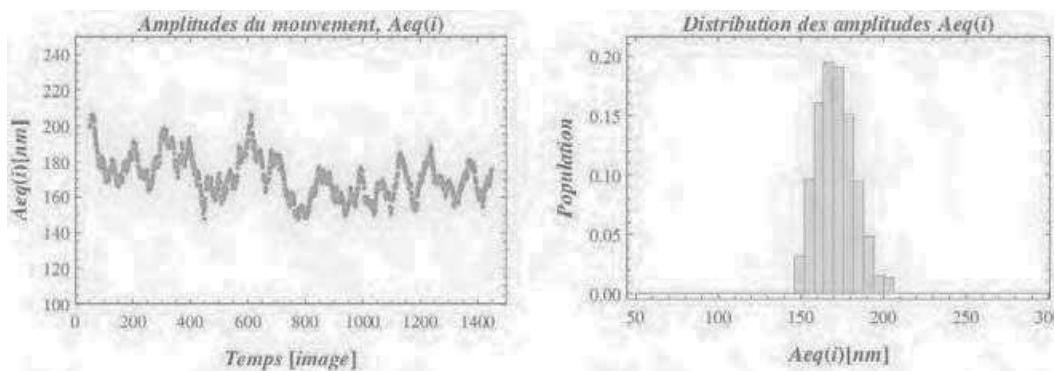


Fig. 3.22 – ① Le tracé de l' **amplitude du mouvement** au cours du temps pour un complexe ADN/particule. ② Les **distributions** des amplitudes du mouvement (ADN  $L_{ADN}=1201$  bp,  $R_{par} = 150$  nm)

Il est possible de construire la distribution de probabilité de  $A_{eq,i}$  de cette trajectoire, Fig. 3.22 ②. Cette distribution est nettement centrée. La moyenne de cette distribution correspond à l'estimation de la **distance bout-à-bout** de notre complexe ADN/particule.

## 3.6 FILTRES DE SÉLECTION APPLIQUÉS PAR LE LOGICIEL *NanoMultiplex*

Sur l'ensemble des trajectoires brutes toutes ne sont pas valides. En effet certaines présentent un comportement erratique. Il est donc nécessaire d'appliquer un certain nombre de filtres afin de séparer les trajectoires valides de l'ensemble des trajectoire brutes.

Les systèmes présentant un comportement très erratique sont directement exclus de la boucle de traitement des données, réalisée par le logiciel *NanoMultiplex*. Cela permet un gain de temps et d'espace de stockage lors du traitement, de l'analyse et de l'écriture des fichiers de données de chaque complexe ADN/particule.

Les deux filtres suivants, sont réalisés sur une phase de pré-mesure. La durée de cette dernière est fixée par nos soins, (*commande : Validation duration (s)*). Le temps minimum de cette étape est, là aussi, estimé à 2 s, afin que chaque système ait le temps de relaxer et que nous obtenions ainsi un maximum d'informations sur le comportement des objets.

### 3.6.1 Amplitude du point d'ancrage

Le calcul de l'**amplitude du point d'ancrage** fournit une bonne information sur la possibilité de collage non spécifique de la bille sur la surface, ou encore de son décrochage au cours du temps. Ce type d'événement, illustré en Fig. 3.23 , induit une brusque variation de l'amplitude du mouvement du point d'ancrage. Tout objet dont le point d'ancrage subit un brusque déplacement peut ne pas être pris en compte dans le reste du traitement des données. Autrement dit, ces objets ne seront pas pris en compte dans l'analyse des données sur la durée complète du temps d'acquisition

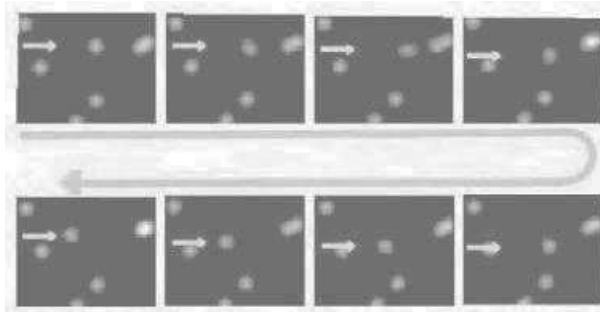


Fig. 3.23 – Illustration en séquence d'image, d'un décrochage de la particule. (ADN  $L_{ADN}=1201$  bp,  $R_{par} = 150$  nm)

Le calcul de l'amplitude du point d'ancrage suit la même expression qu'en Eq. 3.10, sauf qu'elle s'applique sur les  $X_{anc,i}$  et  $Y_{anc,i}$ , et non les  $X_{cor,i}$  et  $Y_{cor,i}$ .

### 3.6.2 Facteur de symétrie

Le **facteur de symétrie** caractérise l'homogénéité de la distribution des positions des billes, Fig. 3.24. Un cercle possède un facteur de symétrie égal à 1, tandis qu'une ellipse aura un facteur de symétrie  $> 1$ . Ce facteur de forme permet de quantifier l'étalement des distributions 2D des  $X_{cor,i}$  et  $Y_{cor,i}$ , et ainsi d'extraire les distributions trop dissymétriques.

En effet, une distribution dissymétrique peut être la signature d'un défaut de réalisation de nos complexes ADN/particule. Si deux ADN sont fixés sur la même nanoparticule son mouvement brownien est fortement dissymétrique en plus de disposer d'une amplitude du mouvement réduite.

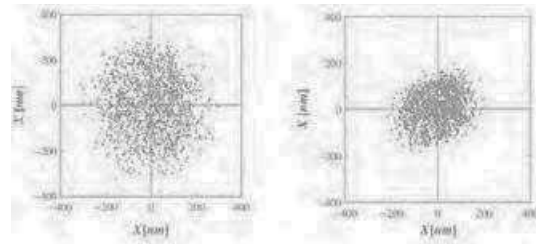


Fig. 3.24 – Illustration de la différence d'asymétrie entre distributions (ADN  $L_{ADN}=1201$  bp,  $R_{par} = 150$  nm)

#### Expression du facteur de symétrie

Le calcul du facteur de symétrie requiert la construction de la matrice de covariance des symétries  $M_{cov} = \begin{vmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{vmatrix}$ . Cette matrice quantifie les écarts conjoints des variables  $X_{cor,i}$  et  $Y_{cor,i}$ , par rapport à leurs espérances respectives.



Les termes de cette matrice s'écrivent donc

$$\begin{aligned}\sigma_{xx} &= \frac{1}{Nbr_{point}} \cdot \sum_{i=1}^{Nbr_{point}} (X_{cor,i})^2 - \frac{1}{Nbr_{point}^2} \cdot \left( \sum_{i=1}^{Nbr_{point}} X_{cor,i} \right)^2 \\ \sigma_{xy} &= \frac{1}{Nbr_{point}} \cdot \sum_{i=1}^{Nbr_{point}} (X_{cor,i} \cdot Y_{cor,i}) - \frac{1}{Nbr_{point}^2} \cdot \left( \sum_{i=1}^{Nbr_{point}} X_{cor,i} \right) \cdot \left( \sum_{i=1}^{Nbr_{point}} Y_{cor,i} \right)\end{aligned}\quad (3.11)$$

La variable  $Nbr_{point}$  est choisie égale à  $Nbr_{image}$ . Ensuite, il faut calculer les valeurs propres de la matrice de covariance,  $\lambda_1$  et  $\lambda_2$ . Enfin, la valeur du facteur de symétrie est définie par le rapport suivant :

$$F_{sym} = \sqrt{\frac{Max(\lambda_1, \lambda_2)}{Min(\lambda_1, \lambda_2)}} \quad \text{Facteur de Symétrie} \quad (3.12)$$

### Valeur seuil du facteur de symétrie

Afin d'éliminer les trajectoires erratiques, nous souhaitons obtenir une estimation du seuil du facteur de symétrie à appliquer sur nos expériences. Pour cela, nous procédons à des simulations Monte-Carlo.

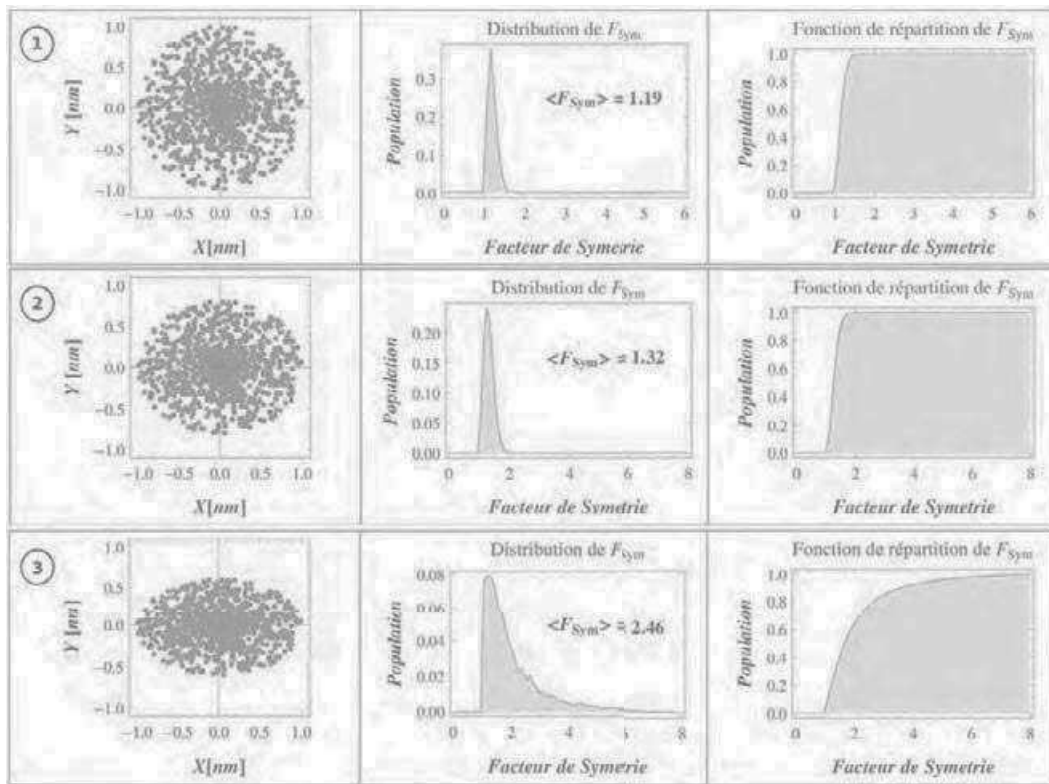


Fig. 3.25 – ① Tirage aléatoire dans un cercle parfait avec une distribution 2D gaussienne, Rayon = (1,1). ② Tirage aléatoire dans une ellipse de Rayon = (1,0.8). ③ Tirage aléatoire dans une ellipse de Rayon = (1,0.6). Tirage de 500000 points.

Pour cela nous tirons aléatoirement des positions dans un cercle parfait, puis nous testons l'influence de différents paramètres d'ellipticité sur le facteur de symétrie de la distribution 2D, Fig. 3.25. Le facteur de symétrie est calculé sur une fenêtre glissante de  $Nbr_{point} = Nbr_{image} = 51$ , comme lors du traitement de nos données expérimentales.

Les mesures expérimentales ne sont pas statistiquement parfaites, il est donc nécessaire de prendre une valeur seuil pour le facteur de symétrie supérieur à la valeur moyenne obtenue pour un cercle parfait où  $\langle F_{Sym} \rangle = 1.25$ . La valeur seuil du facteur de symétrie est prise égale à  $\langle F_{Sym} \rangle_{seuil} = 1.35$ . Nous conservons les trajectoires, jusqu'à celle possédant un caractère d'ellipticité de 0.80. De plus, tout objet dont le facteur de symétrie, moyenné sur la phase de pré-mesure de 2 s, est  $F_{Sym} > \langle F_{Sym} \rangle_{seuil} = 1.35$  n'est pas pris en compte dans le reste du traitement des données.

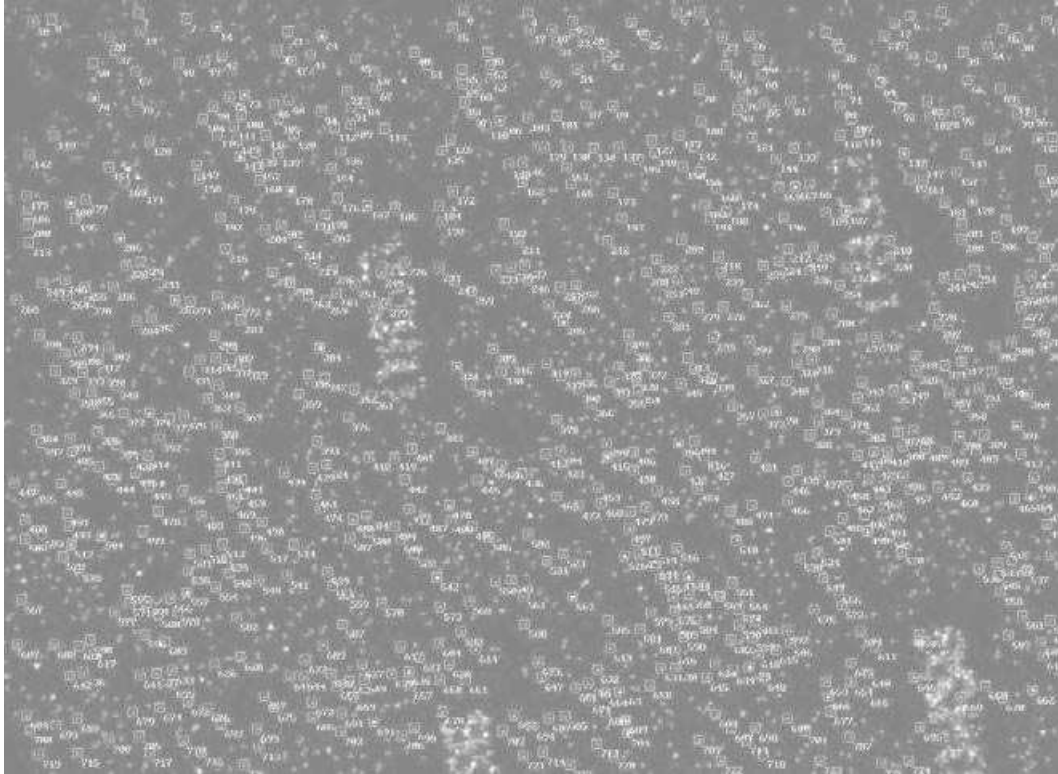


Fig. 3.26 – Champ d'observation en champ sombre, soit  $215 \times 160 \mu\text{m}$ . Les complexes ADN/particule sont réalisés avec  $L_{DNA} = 1201 \text{ bp}$ ,  $R_{par} = 150 \text{ nm}$ , dans des conditions salines proches des conditions physiologiques. Chaque trajectoire considérée comme valide lors de la phase de pré-mesure présente sa zone de R.O.I. visible. Chaque boîte est associée à un numéro qui correspond au numéro d'indexation de son fichier de données.

L'ensemble des complexes ADN/particule remplissant ces deux premiers critères de sélection, sont considérés comme valides après cette phase de pré-mesure. Ainsi ces systèmes sont suivis, leurs positions ainsi que leurs amplitudes du mouvement sont détectées et collectées dans les fichiers de données. Prenons comme exemple le champ d'observation présenté en Fig. 3.17, suite à cette phase de pré-mesure, 727 trajectoires sont suivies comme visualisé en Fig. 3.26.

Toutes les trajectoires ne sont pas obligatoirement valides sur l'ensemble du temps d'acquisition. Il est donc nécessaire d'ajouter des filtres de sélectivité caractérisant le comportement du système sur l'ensemble de la trajectoire et non pas seulement sur la phase de pré-mesure.

De plus, nous cherchons à remonter aux propriétés physiques et mécaniques de l'ADN. Jusqu'à présent, la grandeur mesurée correspond à l'amplitude du mouvement de l'ensemble du complexe ADN/particule, nous devons donc en extraire la contribution de l'ADN. Pour cela nous devons être extrêmement fin sur notre mesure. Celle-ci doit être corrigée des erreurs, et des biais engendrés par nos appareils de mesure eux-mêmes. Les erreurs de pointé ou les effets engendrés par le moyennage vidéo des positions durant le temps d'acquisition doivent être examinés.

À cet égard, durant ma thèse nous avons élaboré un protocole de traitement et d'analyse de données. Celui-ci est codé via le logiciel *Mathematica*, l'ensemble du code de ce script est en Annexe C. Ci-dessous, les étapes de traitement et d'analyse des données applicables pour un **système dans un seul état**, sont présentées

### 3.7 CORRECTIONS ET PREMIÈRES ÉTAPES DE L'ANALYSE DE DONNÉES

Tout d'abord il est primordial d'insister sur la nature de l'état dans lequel se trouve le système étudié. En effet, sauf mention expresse du contraire, afin de réaliser nos mesures expérimentales nous fixons le système dans un état donné. Ainsi, les paramètres mesurés caractérisent cet état et non une transition d'état. De ce fait, la construction de la distribution de probabilité des amplitudes du mouvement du système est caractérisée par une population et une seule. C'est dans ce cadre-là que la procédure d'analyse ci-dessous est développée, c'est-à-dire une seule population, ou un seul état, définit nos systèmes ADN/particules. Pour l'instant, nous ne nous intéressons donc pas à proprement parler à la dynamique d'événements intervenant sur le système. Notons, qu'il est possible d'activer des balises qui suppriment l'activation de certains filtres, détaillés ci-dessous, afin de ne pas supprimer des sous-populations qui définissent la physique du système.

Ensuite, il est nécessaire de faire la distinction entre certaines variables que nous allons utiliser tout au long du traitement des données. Nous parlerons de la valeur moyenne de l'amplitude du mouvement en terme de **moyenne mathématique** et en terme de **moyenne extraite via un ajustement gaussien** de la distribution de ces amplitudes.

Pour une trajectoire d'une minute, le nombre de points construisant l'amplitude du mouvement au cours du temps  $A_{eq,i}$  est égal à  $Nbr_{image} = T_{acq.} \cdot f_{acq.} = 51 \times 60 \times 25 = 1500$ . Nous pouvons extraire la **moyenne mathématique** de l'amplitude du mouvement que nous noterons  $A_{eq,[MOY]}$ .

De plus, nous pouvons construire la distribution des valeurs  $A_{eq,i}$ . Une autre estimation de la valeur moyenne de la distribution des  $A_{eq,i}$  peut se faire via un **ajustement gaussien** de cette distribution. La moyenne  $\mu$  de cette gaussienne est notée  $\mu_{A_{eq,i}}$ . Comme nous le verrons par la suite, les valeurs extraites via la **moyenne mathématique** ou via un **ajustement gaussien** sont proches mais pas exactement égales.

$$A_{eq,i} \begin{cases} \xrightarrow[\text{sur } i]{\text{Moyenne}} A_{eq,[Moy]} \\ \xrightarrow[\text{sur } i]{\text{Fit G.}} \mu_{eq,i} \end{cases} \begin{matrix} \# = 1 \\ \# = 1 \end{matrix}$$

Maintenant, revenons à l'ensemble de nos trajectoires. Nous noterons  $Nbr_{trajectoire} = j$  le nombre et le numéro des trajectoires. Nous pouvons les analyser de la même façon. Chaque trajectoire  $j$ , est associée à son amplitude du mouvement  $A_{eq,i,j}$  au cours du temps. La **moyenne mathématique** de son amplitude du mouvement devient  $A_{eq,[MOY],j}$ . Nous pouvons construire une distribution de ces moyennes. Cette dernière sera donc constituée de  $j$  valeurs statistiquement indépendantes correspondant aux  $A_{eq,[MOY],j}$ .

Cette distribution possède aussi une **moyenne mathématique**,  $\langle A_{eq,[MOY],[MOY]} \rangle$ , la *moyenne des moyennes*. De plus, cette distribution possède aussi une estimation de sa moyenne lorsqu'elle est ajustée par une gaussienne. Soit  $\langle \mu_{A_{eq,[MOY]}} \rangle$  la **moyenne gaussienne** de la distribution des moyennes des distributions.

$$A_{eq,i,j} \begin{cases} \xrightarrow[\text{sur } i]{\text{Moyenne}} A_{eq,[Moy],j} \\ \xrightarrow[\text{sur } i]{\text{Fit G.}} \mu_{eq,i,j} \end{cases} \begin{matrix} \# = 1 * Nbr_{trajectoire} \\ \# = 1 \end{matrix} \begin{cases} \xrightarrow[\text{sur } j]{\text{Moyenne}} \langle A_{eq,[Moy],[Moy]} \rangle \\ \xrightarrow[\text{sur } j]{\text{Fit G.}} \langle \mu_{eq,[Moy],j} \rangle \end{cases} \begin{matrix} \# = 1 \\ \# = 1 \end{matrix}$$

De plus, là encore nous pouvons construire la distribution de l'ensemble des  $Nbr_{image}$  pour l'ensemble des  $Nbr_{trajectoire}$ . Autrement dit, cette distribution sera constituée de  $Nbr_{image} \times Nbr_{trajectoire}$  points. La **moyenne mathématique** de cette distribution sera notée  $\langle A_{eq,i,j} \rangle$ . Tandis que, la valeur moyenne de l'**ajustement gaussien** de cette dernière distribution se notera  $\langle \mu_{A_{eq,i,j}} \rangle$ .

$$A_{eq,i,j} \begin{cases} \xrightarrow[\text{sur } i]{\text{Moyenne}} A_{eq,[Moy],j} \\ \xrightarrow[\text{sur } i]{\text{Fit G.}} \mu_{eq,i,j} \end{cases} \begin{matrix} \# = 1 * Nbr_{trajectoire} \\ \# = 1 \end{matrix} \begin{cases} \xrightarrow[\text{sur } j]{\text{Moyenne}} \langle A_{eq,[Moy],[Moy]} \rangle \\ \xrightarrow[\text{sur } j]{\text{Fit G.}} \langle \mu_{eq,[Moy],j} \rangle \end{cases} \begin{matrix} \# = 1 \\ \# = 1 \end{matrix}$$

$$A_{eq,i,j} \begin{cases} \xrightarrow[\text{sur } i, j]{\text{Moyenne}} \langle A_{eq,i,j} \rangle \\ \xrightarrow[\text{sur } i, j]{\text{Fit G.}} \langle \mu_{eq,i,j} \rangle \end{cases} \begin{matrix} \# = 1 \\ \# = 1 \end{matrix}$$

En résumé, il nous est possible de travailler sur les moyennes de chacune des distributions ou de travailler sur l'ensemble des points de chacune des distributions. En fonction de l'étape de traitement des données, nous travaillerons soit avec les valeurs obtenues par le calcul de la **moyenne mathématique**, soit avec les valeurs obtenues par un **ajustement gaussien**.

### 3.7.1 Filtre sur le facteur de symétrie, $F_{Sym}$

Nous calculons la moyenne du facteur de symétrie sur l'ensemble des points,  $i$ , de la trajectoire. La valeur  $\langle F_{Sym} \rangle$  doit être inférieure ou égale au seuil de 1.35 pour que la trajectoire soit conservée. Le facteur de symétrie est calculé pour chacune des trajectoires. Les trajectoires ne respectant pas ce critère sont rejetées.

Les trajectoires conservées remplissent le critère de sélectivité suivant :

$$\langle F_{Sym} \rangle \geq 1.35 \quad \text{Filtre : } F_{Sym} \quad (3.13)$$

Pour illustrer notre propos, continuons l'analyse des données obtenues sur les complexes ADN/particule réalisés avec  $L_{ADN} = 1201 \text{ bp}$ ,  $R_{par} = 150 \text{ nm}$ . L'exemple en Fig. 3.26, représente un jour d'expérience. Étoffons nos statistiques en y ajoutant deux autres journées d'expériences supplémentaires, réalisées sur le même système et dans les mêmes conditions.

Le nombre de fichiers de données obtenus sur ces 3 jours est de 2285 trajectoires. Suite à l'application du filtre sur le facteur de symétrie ci-dessus, 1118/2285 trajectoires sont conservées.

### 3.7.2 Filtre sur les amplitudes, $A_{eq,i,j}$

Pour l'application des deux filtres suivants, nous testons chaque valeur de l'amplitude au cours du temps et ce, pour chacune des trajectoires. Ce sont donc toutes les valeurs  $A_{eq,i,j}$  qui sont testées.

Toute trajectoire possédant au moins une valeur d'amplitude au cours du temps ne satisfaisant par les deux conditions ci-dessous est rejetée. Car, dans cette configuration nous considérons que l'ensemble de la trajectoire est erratique.

#### Exclusion des nanoparticules se collant à la surface

Si au cours du temps une nanoparticule a interagi non spécifiquement avec la surface, nous rejetons l'ensemble de sa trajectoire. Ce type d'événement est caractérisé par une valeur d'amplitude qui passe brusquement en dessous de 1 nm. Cette très faible valeur d'amplitude est liée à la perte de focus sur la nanoparticule, due à sa brusque variation de position suivant l'axe Z.

Les trajectoires conservées remplissent le critère de sélectivité suivant :

$$A_{eq,i,j} > 1 \text{ nm} \quad \text{Filtre : } A_{eq,i,j} \text{ ①} \quad (3.14)$$

#### Exclusion des nanoparticules se décrochant de la surface

Si une nanoparticule se décroche de la surface au cours du temps<sup>6</sup>, nous rejetons l'ensemble de sa trajectoire. Ce type d'événement est caractérisé par une valeur d'amplitude qui passe brusquement en dessous de  $1 \times 10^{-3} \text{ nm}$ , avant la

6. Pour ce type d'événement nous ne pouvons savoir si c'est la nanoparticule qui se décroche de l'ADN, si c'est l'ensemble du complexe ADN/particule, ou même si c'est le complexe plus la streptavidine qui se décroche de la surface.

perte du signal. Lors de l'écriture des fichiers le logiciel *NanoMultiplex* indexe ce type de valeurs d'amplitude du mouvement sous le sigle *IND* correspondant au terme *indéfini*. Toute trajectoire possédant au moins une *valeur* d'amplitude mise à *IND* est rejetée.

Les trajectoires conservées remplissent le critère de sélectivité suivant :

$$A_{eq,i,j} \neq IND \quad \text{Filtre : } A_{eq,i,j} \text{ ②} \quad (3.15)$$

Suite à l'application de ces deux filtres sur les valeurs des amplitudes du mouvement ci-dessus, 1101/1118 trajectoires sont conservées.

### 3.7.3 Distribution des $A_{eq,[Moy],j}$

#### Construction de la distribution des $A_{eq,[Moy],j}$

Nous construisons la distribution de probabilité constituée de l'ensemble des moyennes mathématiques de l'amplitude du mouvement au cours du temps de chaque trajectoire  $A_{eq,[Moy],j}$ , Fig. 3.27. Cette distribution est donc constituée de  $Nbr_{trajectoire}$  variable aléatoire.

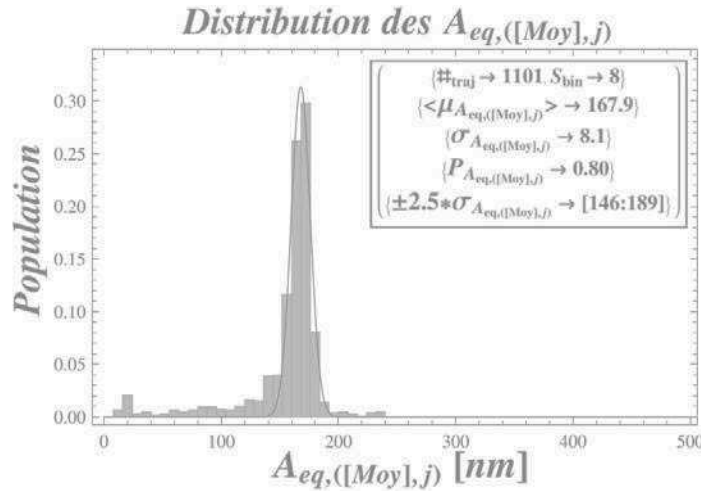


Fig. 3.27 – Distribution de probabilité des  $A_{eq,[Moy],j}$ , ( $ADN L_{ADN} = 1201 \text{ bp}$ ,  $R_{par} = 150 \text{ nm}$ )

La taille des *bins* de l'histogramme est définie telle que :

$$S_{bin, \text{①}} = \frac{Max(A_{eq,[Moy],j}) - Min(A_{eq,[Moy],j})}{\sqrt{Nbr_{trajectoire}}} \quad \text{Taille bin ①} \quad (3.16)$$

La distribution est estimée par une densité de probabilité gaussienne.

$$P(x) = S_{bin, \text{①}} \times P_{op, \text{①}} \cdot \left( \frac{\exp\left[-\frac{(x-\mu_{\text{①}})^2}{2\sigma_{\text{①}}^2}\right]}{\sigma_{\text{①}} \sqrt{2\pi}} \right) \quad \text{Fit Gaussien} \quad (3.17)$$

Cet ajustement nous permet d'extraire :

- $\mu_{\textcircled{1}} \Rightarrow \langle \mu_{Aeq,([Moy],j)} \rangle_{\textcircled{1}}$
- $\sigma_{\textcircled{1}} \Rightarrow \sigma_{Aeq,([Moy],j)}_{\textcircled{1}}$

Certaines des valeurs, à faible poids statistique, sont positionnées de part et d'autre des *pieds* de la gaussienne. Dans cette mesure expérimentale HT-TPM seule une longueur d'ADN est présente ainsi qu'une seule taille de nanoparticule. De plus, les conditions de température et de salinité sont fixes et aucune molécule n'est ajoutée afin d'interagir avec notre système. Du fait même, de la nature de cette expérience nous ne pouvons aboutir à l'observation de plusieurs populations possédant un sens physique cohérent et pertinent. En effet, nous nous attendons à un comportement uniforme et semblable de tous les complexes ADN/particule, engendrant une seule population finale. Il est donc acceptable d'associer un caractère erratique à ces valeurs différenciables. Nous allons donc ajouter un filtre de sélectivité afin de différencier et rejeter ces trajectoires.

#### Filtre sur les $A_{eq,([Moy],j)}$

Ce filtre est basé sur les valeurs extraites à la suite de l'ajustement de la distribution par une gaussienne. Là encore, nous travaillons sur l'ensemble des moyennes mathématiques de l'amplitude du mouvement au cours du temps de chaque trajectoire  $A_{eq,([Moy],j)}$ . Toute trajectoire ne satisfaisant pas ce critère sera rejetée

Les trajectoires conservées remplissent le critère de sélectivité suivant :

$$\mu_{\textcircled{1}} - 2.5 \times \sigma_{\textcircled{1}} \leq A_{eq,([Moy],j)} \leq \mu_{\textcircled{1}} + 2.5 \times \sigma_{\textcircled{1}} \quad \text{Filtre : } A_{eq,([Moy],j)}_{\textcircled{1}} \quad (3.18)$$

Rejeter  $(\mu_{\textcircled{1}} \pm 2.5 \times \sigma_{\textcircled{1}})$  correspond à soustraire moins de 2% de la population de la distribution contenue dans l'ajustement gaussien. La suppression de quelques trajectoires complètes rend la nouvelle distribution de probabilité *propre*, Fig. 3.28.

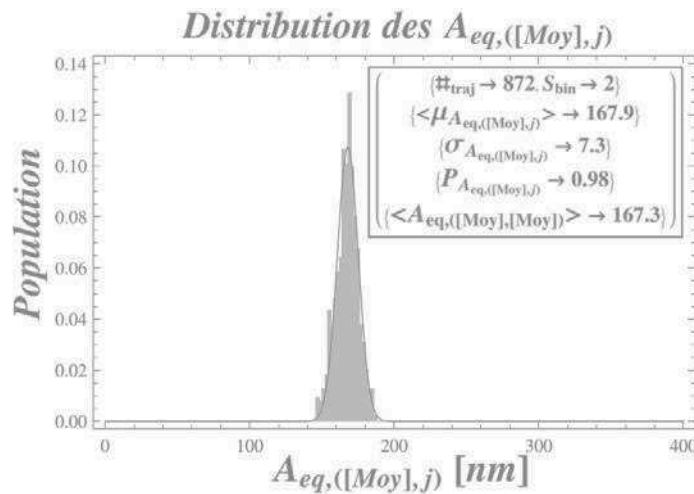


Fig. 3.28 – Nouvelle distribution de probabilité des  $A_{eq,([Moy],j)}$ , (ADN  $L_{ADN} = 1201$  bp,  $R_{par} = 150$  nm)

L'ajustement gaussien est maintenant parfaitement centré sur les amplitudes du mouvement, n'étant plus perturbé par les quelques trajectoires erratiques précédemment observées. Ceci signifie que le comportement que nous venons de filtrer était associé à seulement certaines trajectoires, et non à certains points au cours du temps sur l'ensemble des trajectoires. En effet, si ce comportement était commun à l'ensemble de la population malgré la suppression des quelques trajectoires contribuant principalement à ce phénomène, l'étalement de la distribution resterait visible mais dans un pourcentage plus faible. Le comportement que nous venons d'exclure correspond bien à un comportement minoritaire de seulement quelques individus.

La taille des *bins* est à nouveau calculée, avec la même équation que précédemment. Leur valeur,  $S_{bin, ②}$ , est réduite par rapport à  $S_{bin, ①}$ . Ceci est dû à la réduction de l'étalement de la distribution.

Cet ajustement nous permet d'extraire :

$$\begin{aligned} - \mu_{②} &\Rightarrow \langle \mu_{Aeq, [Moy], j} \rangle_{②} \\ - \sigma_{②} &\Rightarrow \sigma_{Aeq, [Moy], j} \text{ ②} \end{aligned}$$

Suite à l'application de ce filtre sur les valeurs des moyennes mathématiques de chaque amplitude du mouvement ci-dessus, 872/1101 trajectoires sont conservées.

La distribution ci-dessous, Fig. 3.29, correspond à l'ensemble des amplitudes du mouvement au cours du temps pour chaque trajectoire. Tous les points d'amplitude du mouvement des trajectoires sélectionnées y sont contenus.

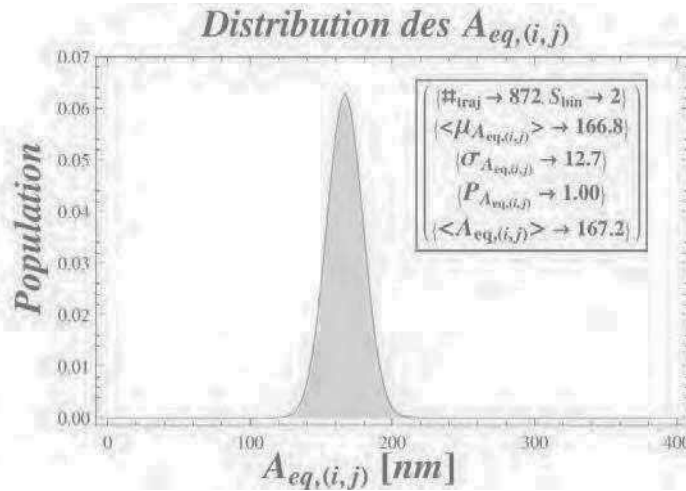


Fig. 3.29 – Distribution de probabilité de tous les points  $A_{eq,i,j}$ . (ADN  $L_{ADN} = 1201$  bp,  $R_{par} = 150$  nm)

Ce nouvel ajustement sur l'ensemble des points, nous permet d'extraire :

$$\begin{aligned} - \mu_{③} &\Rightarrow \langle \mu_{Aeq,i,j} \rangle_{①} \\ - \sigma_{③} &\Rightarrow \sigma_{Aeq,i,j} \text{ ①} \end{aligned}$$



### 3.7.4 Valeur brute de l'amplitude du mouvement

À la fin de ces étapes, la valeur **brute** de l'amplitude du mouvement de nos complexes ADN/particule que nous retenons correspond à la moyenne extraite via l'ajustement de la distribution de tous les  $A_{eq,i,j}$  par une gaussienne. Cette amplitude est associée à la dénomination **brute** car aucune correction des biais engendrés par nos appareils de mesure n'est appliquée.s

### 3.7.5 Correction de l'effet du moyennage vidéo

Dans nos mesures TPM, les positions extraites sur une image  $i$  ne correspondent pas à une seule position du complexe DNA/particule à un instant  $t$ . Une image est une moyenne des positions explorées durant le temps d'exposition  $T_{ex}$ .

L'intensité de ce phénomène dépend du rapport entre le **temps de corrélation**,  $\tau_{corr}$ , du système ADN/particule et le **temps d'exposition**,  $T_{ex}$ , de l'appareil de mesure. Si le temps caractéristique du système n'est pas trop faible par rapport  $T_{ex}$ , cet effet peut être corrigé [DS06].

Nous pouvons corriger l'effet de moyennage vidéo selon les limites suivantes :

$$\begin{aligned} \tau_{corr} \geq \frac{2.T_{ex}}{3} &\Rightarrow \text{Correction} \\ \tau_{corr} < \frac{2.T_{ex}}{3} &\Rightarrow \text{Pas de correction} \quad \text{Limite correction } \tau_{corr} \end{aligned} \quad (3.19)$$

Plusieurs étapes de calculs sont nécessaires afin de corriger cet effet. En premier lieu, nous devons calculer les différents termes de la fonction de corrélation pour chaque trajectoire. Ensuite, le rapport des fonctions lorsque les positions sont en phase ou hors phase, nous permet d'extraire le **temps de corrélation**,  $\tau_{corr}$ , des systèmes. Il est alors facile de remonter au **temps de relaxation**,  $\tau_{||}$ , des systèmes. Enfin, les valeurs **brutes** des amplitudes du mouvement peuvent être corrigées.

#### Fonction de corrélation des positions en phase $C_m(0)$

Nous mesurons la corrélation entre les valeurs des positions extraites via la mesure de  $A_{eq}$  au temps  $t = i$  par rapport à cette dernière au temps  $t = i + s$ . Le calcul de la fonction de corrélation lorsque les positions sont en phase  $C_m(s = 0)$ , autrement dit cela correspond à la comparaison de  $A_{eq}(t = i)$  par rapport à elle-même, nous sert de fonction de référence. Son expression est la suivante :

$$C_m(s = 0) = \langle A_{eq}(t = i + 0).A_{eq,i}(t = i + 0) \rangle \quad (3.20)$$

Selon les positions  $X_{corr}$  et  $Y_{corr}$ , l'expression directe de la fonction de corrélation est :

$$C_m(s = 0) = \frac{1}{N} \sum_{i=1}^N (X_{corr}^2(t = i) + Y_{corr}^2(t = i)) \quad (3.21)$$

$$(3.22)$$

### L'erreur de pointé

La position détectée par notre dispositif expérimental correspond à la somme de la position *réelle* et de l'**erreur de pointé**,  $\Delta err$ , existant entre deux variables aléatoires indépendantes. Les valeurs *réelles* des positions sont donc égales à  $X_{corr}^0 = X_{corr} - \Delta err_X$ . La valeur *réelle* de l'amplitude du mouvement est donc égale à  $\sqrt{\langle A_{eq}^2 \rangle - \Delta^2 err}$ .

L'**erreur de pointé** correspond à l'incertitude sur la mesure des positions. Autrement dit, elle correspond à l'amplitude du mouvement des complexes immobiles fixés sur la surface. L'estimation de la valeur de  $\Delta err$  est simple à estimer. Il suffit d'étudier la distribution des positions d'un objet fixe et d'extraire l'estimation de son amplitude du mouvement, Fig. 3.30.

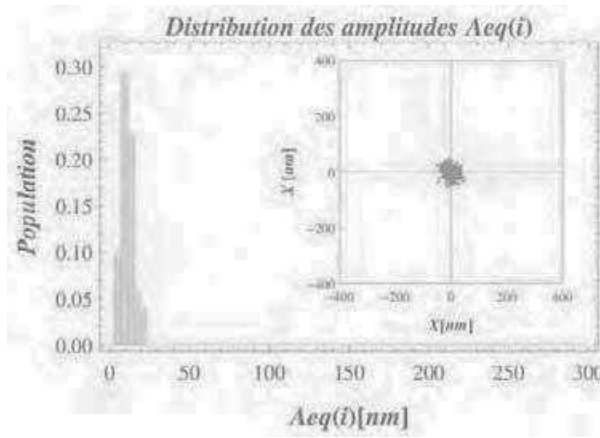


Fig. 3.30 – La distribution de probabilité et dans l'encadré les *positions relatives*. ( $ADN L_{ADN} = 1201 \text{ bp}$ ,  $R_{par} = 150 \text{ nm}$ )

L'erreur de pointé est donc égale à la moyenne extraite par l'application d'un ajustement gaussien sur les estimations des amplitudes du mouvement d'objet fixe. L'estimation de cette moyenne sur un échantillon d'une centaine de systèmes fixes donne une erreur de pointé égale à ;

$$\Delta err_{pt} = \langle \mu A_{eq,i,j} \rangle_{(Particule\ fixe)} = 10 \text{ nm} \quad \Delta err$$

Il est possible d'appliquer la correction liée à l'**erreur de pointé**,  $\Delta err$ , directement sur le calcul de la fonction de corrélation des positions en phase  $C_m(s=0)$ .

$$C_m(s=0)_{err} = C_m(s=0) - \Delta^2 err \quad (3.23)$$

$$\Rightarrow = \frac{1}{N} \sum_{i=1}^N (X_{corr}^2(t=i) + Y_{corr}^2(t=i)) - \Delta^2 err \quad (3.24)$$

Afin de ne pas alourdir le détail du calcul, nous nous intéressons à la fonction de corrélation des positions  $X_{corr}$  à laquelle est associée l'erreur de pointé  $\Delta err_X$ . Le calcul de cette dernière suit le détail suivant :

$$C_{m,X}(s)_{err} = \langle X_{corr}^0(t) \cdot X_{corr}^0(t+s) \rangle \quad (3.25)$$

$$\Rightarrow = \langle [X_{corr}(t) - \Delta err_X(t)] \cdot [X_{corr}(t+s) - \Delta err_X(t+s)] \rangle \quad (3.26)$$

L'expression de la fonction de corrélation des positions en phase,  $s = 0 \Rightarrow t = i$ , s'exprime donc tel que :

$$\Rightarrow C_{m,X}(s=0)_{err} = \langle [X_{corr}(t=i) - \Delta err_X(t=i)]^2 \rangle \quad (3.27)$$

$$\Rightarrow = \frac{1}{N} \sum_{i=1}^N X_{corr}^2(i) - \Delta^2 err_X \quad (3.28)$$

### Fonction de corrélation des positions hors phase $C_m(s)$

Le détail de l'expression de la fonction de corrélation,  $C_m(s)$ , selon le décalage successif,  $s$ , des positions  $X_{corr}$  s'exprime :

$$\begin{aligned} C_{m,X}(s)_{err} &= \langle [X_{corr}(t) - \Delta err_X(t)] \cdot [X_{corr}(t+s) - \Delta err_X(t+s)] \rangle \\ \Rightarrow &= \langle X_{corr}(t) \cdot X_{corr}(t+s) \rangle + \langle X_{corr}(t) \cdot \Delta err_X(t+s) \rangle \\ \Rightarrow &+ \langle X_{corr}(t+s) \cdot \Delta err_X(t) \rangle + \langle \Delta err_X(t+s) \cdot \Delta err_X(t) \rangle \end{aligned} \quad (3.29)$$

Les trois derniers termes de cette équation sont nuls. En effet, nous considérons que toutes les sources de bruits, sauf l'erreur induite par la dérivée de la plaine du microscope, sont décorrélées pour tous temps  $s' \neq s$ . Il en découle que la prise en compte de l'erreur de pointé se fait seulement via le terme  $C_m(0)_{err}$ .

L'expression de la fonction de corrélation,  $C_m(s)$ , selon le décalage successif,  $s$ , des positions  $X_{corr}$  et  $Y_{corr}$  s'exprime :

$$C_m(s) = \frac{1}{N-s} \sum_{i=1}^{N-s} (X_{corr}(i) \cdot X_{corr}(i+s) + Y_{corr}(i) \cdot Y_{corr}(i+s)) \quad (3.30)$$

La fonction ainsi construite possède des points espacés tous les  $s = n \times T_{ac}$ .

### Rapport des fonctions de corrélation

Ainsi, pour chaque trajectoire TPM nous calculons l'ensemble de ses fonctions de corrélation. Puis, le rapport  $C_m(s)/C_m(0)_{err}$  est calculé, Fig. 3.31.

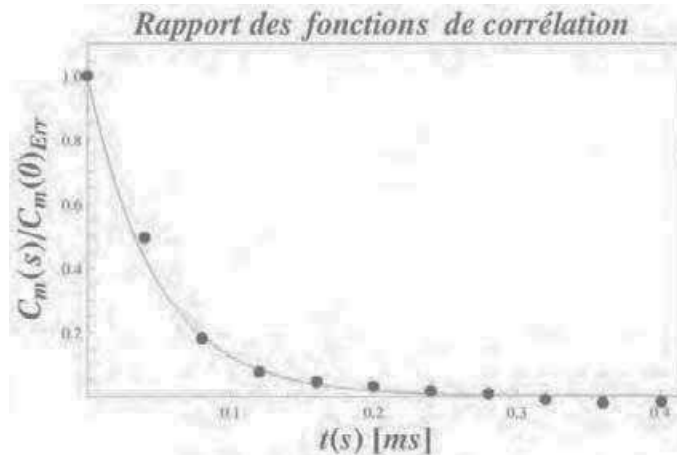


Fig. 3.31 – Expression et fit de  $C_m(s)/C_m(0)_{err}$ , ( $L_{ADN} = 1201 \text{ bp}$ ,  $R_{par} = 150 \text{ nm}$ )

### Extraction du temps de corrélation

L'ajustement du rapport  $C_m(t+s)/C_m(0)_{err}$  selon l'équation ci-dessous permet d'extraire le **temps de corrélation**  $\tau_m$  qui va nous permettre de calculer le **temps de relaxation**  $\tau_{||}$  associé à chaque trajectoire TPM, [MTB<sup>+</sup>10].

$$\frac{C_m(s)}{C_m(0)_{err}} = (1 + 2 \frac{\tau_{corr}}{T_{av}}) \exp[-\frac{t}{\tau_{corr}}] - 2 \frac{\tau_{corr}}{T_{av}} \quad (3.31)$$

$$\Rightarrow \tau_{||} = \tau_{corr} - \frac{T_{ex}}{3} \quad \text{Temps de relaxation} \quad (3.32)$$

Pour affiner la précision de cette mesure, trois ajustements successifs sont réalisés sur une gamme dont l'étendue est progressivement réajustée. Le premier ajustement donnant accès à  $\tau_{corr,①} \Rightarrow \tau_{||,①}$  est calculé sur l'ensemble du domaine des  $t(s)$ . Le deuxième est réalisé sur le domaine restreint  $Dom② : [0 : 4 \times \tau_{||,①}]$ , il donne accès à  $\tau_{||,②}$ . De, même pour le troisième ajustement,  $Dom③ : [0 : 4 \times \tau_{||,②}] \Rightarrow \tau_{||,③}$ . La valeur de  $\tau_{||,③}$  est la valeur finale associée à  $\tau_{||}$  pour cette trajectoire.

### 3.7.6 Correction de l'effet du moyennage vidéo

Les amplitudes **brutes**,  $A_{eq,i,j}$ , d'une trajectoire  $j$  sont corrigées de l'effet de moyennage vidéo, dépendant de son **temps de relaxation**  $\tau_{||,j}$ , selon l'équation suivante, [MTB<sup>+</sup>10] :

$$R_{||,i,j} = A_{eq,i,j} \cdot [2 \frac{\tau_{||,j}}{T_{ex}} - 2 (\frac{\tau_{||,j}}{T_{ex}})^2 (1 - \exp[-\frac{T_{ex}}{\tau_{||,j}}])]^{(1/2)} \quad (3.33)$$

Ainsi, chaque valeur d'amplitude du mouvement est corrigée en fonction de son temps de corrélation.

### Distribution des $R_{||,i,j}$

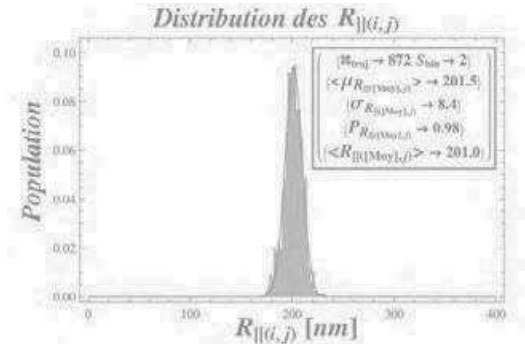


Fig. 3.32 – Distribution des moyennes mathématiques des amplitudes du mouvement corrigées par l'effet du temps de corrélation,  $R_{||,[Moy],j}$ . ( $L_{ADN} = 1201$  bp,  $R_{par} = 150$  nm).

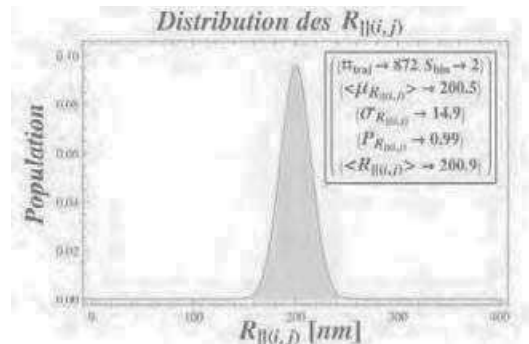


Fig. 3.33 – Distribution de l'ensemble des valeurs des amplitudes du mouvement, le tout corrigé par l'effet du temps de corrélation,  $R_{||,i,j}$ . ( $L_{ADN} = 1201$  bp,  $R_{par} = 150$  nm).

La construction de ces distributions, Fig. 3.32 et 3.33, donnent des valeurs de la moyenne des amplitudes extraites par estimation gaussienne quasiment similaire, soit  $\langle \mu_{R_{||, Moy}, j} \rangle = 201.5 \text{ nm}$  et  $\langle \mu_{R_{||, i, j}} \rangle = 200.5 \text{ nm}$ .

### Distribution des $\tau_{||, j}$

Nous pouvons comparer la représentation du graphique de densité de l'amplitude du mouvement, avant correction ( $A_{eq, i, j}$ ), et après correction ( $R_{||, i, j}$ ) du temps de corrélation. En ①, nous distinguons la corrélation entre ces deux grandeurs. Lorsque  $A_{eq, i, j}$  augmente,  $\tau_{||, j}$  augmente. Après l'application de la correction que nous venons de calculer, ②, ce lien de dépendance diminue. Pour vérifier cela nous calculons la facteur de corrélation de Pearson, grâce à la fonction *mathematica* associée, afin d'estimer le degré de corrélation, autrement dit la dépendance linéaire, entre nos deux variables.

Le facteur de corrélation de Pearson est égal à 1 si l'une des variables est une fonction affine croissante de l'autre variable et, à -1, si la des variable est une fonction affine et décroissante, Fig. 3.34. Plus le coefficient est proche des valeurs extrêmes -1 et 1, plus la corrélation entre les variables est forte. Les valeurs intermédiaires renseignent sur le degré de corrélation, de dépendance linéaire, entre les deux variables, comme nous pouvons en voir l'illustration ci-dessous.

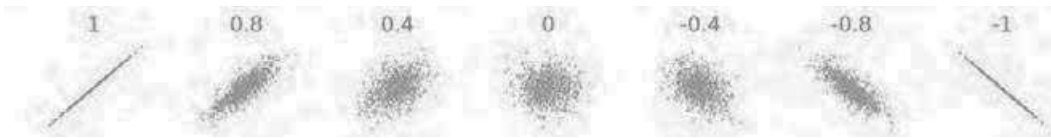


Fig. 3.34 – Exemples de coefficients de corrélation linéaire entre deux variables

L'estimation du facteur de corrélation entre le temps de corrélation et l'amplitude du mouvement est plus proche de zéro, pour les amplitudes corrigées  $R_{||, i, j}$ , Fig. 3.35 ①, donc la corrélation est plus faible, que pour les amplitudes brutes  $A_{eq, i, j}$ , Fig. 3.35 ②.

$L_{DNA}$	Facteur de Pearson	
	$\tau_{  , j} - A_{eq, i, j}$	$\tau_{  , j} - R_{  , i, j}$
1201 bp	0.46	-0.18
575 bp	0.44	-0.21

L'application de la correction diminue la corrélation entre nos deux variables. Cette correction semble *dilater* les valeurs d'amplitude du mouvement  $R_{||, i, j}$ .

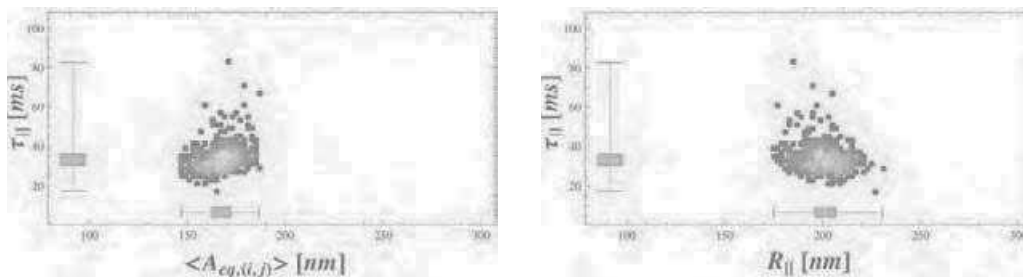


Fig. 3.35 – Graphique de densité de : ① l'amplitude du mouvement  $A_{eq, i, j}$  avant et ② l'amplitude du mouvement  $R_{||, i, j}$  après correction du temps de corrélation. ( $L_{ADN} = 1201 \text{ bp}$ ,  $R_{par} = 150 \text{ nm}$ )

La corrélation entre  $A_{eq,i,j}$  et  $\tau_{||,j}$  est aussi observable sur les molécules d'ADN de petite taille, Fig. 3.36. Rappelons que plus  $L_{ADN}$  est petit, plus le temps de relaxation du système diminue.

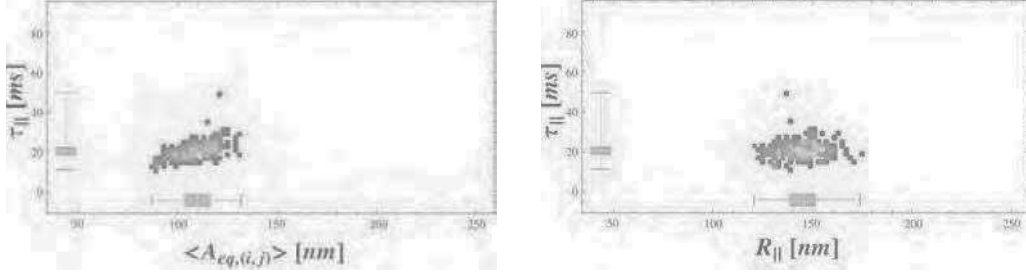


Fig. 3.36 – ( $L_{ADN} = 575$  bp],  $R_{par} = 150$  nm]). Graphique de densité de : ① l'amplitude du mouvement  $A_{eq,i,j}$  avant et ② l'amplitude du mouvement  $R_{||,i,j}$  après correction du temps de corrélation.

Nous nous rapprochons de la limite de correction définie sur  $\tau_{corr}$ , Eq. 3.19.

$$\begin{aligned} \text{Correction ssi } \tau_{corr} \geq \frac{2.T_{ex}}{3} \quad \text{alors } \tau_{||} &= \tau_{corr} - \frac{T_{ex}}{3} \\ \Rightarrow \lim \tau_{||} &= \frac{T_{ex}}{3} = 13 \text{ ms} \quad \tau_{||} \text{ limite} \end{aligned} \quad (3.34)$$

### 3.7.7 Filtre sur $\tau_{||}$

La distribution de l'ensemble des temps de relaxation de nos systèmes, Fig. 3.37, présente quelques trajectoires dont la valeur de  $\tau_{||}$  s'éloigne de la moyenne. Ces trajectoires sont aussi visualisables sur les graphiques de densité. Nous allons donc ajouter un nouveau filtre de sélection suivant la même construction qu'en 3.7.3 .

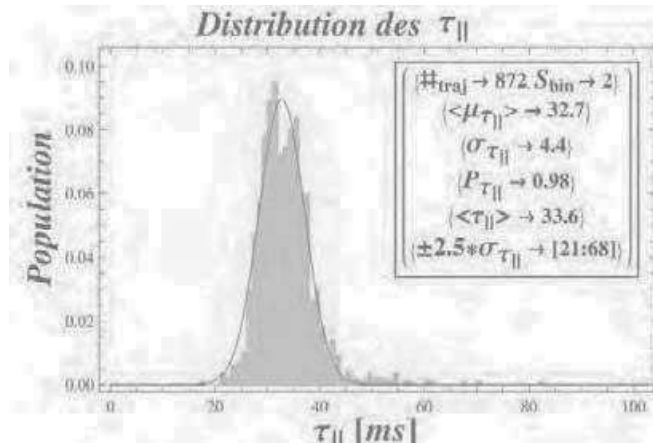


Fig. 3.37 – Distribution des temps de relaxation de l'ensemble des systèmes, ( $L_{ADN} = 1201$  bp,  $R_{par} = 150$  nm)

Un ajustement gaussien appliqué à l'ensemble des  $\tau_{||}$ , nous permet d'extraire les grandeurs suivantes :

- $\mu_{\textcircled{4}} \Rightarrow \langle \mu_{\tau_{||}} \rangle_{\textcircled{1}}$
- $\sigma_{\textcircled{4}} \Rightarrow \sigma_{\tau_{||}}_{\textcircled{1}}$

Toute trajectoire ne satisfaisant pas le critère sur les  $\tau_{||}$ , défini en suivant, sera rejetée. Seules, les trajectoires remplissant le critère de sélectivité suivant sont conservées :

$$\mu_{\textcircled{4}} - 2.5 \times \sigma_{\textcircled{4}} \leq \tau_{||} \leq \mu_{\textcircled{4}} + 2.5 \times \sigma_{\textcircled{4}} \quad \text{Filtre : } \tau_{||} \textcircled{1} \quad (3.35)$$

Suite à l'application de ce filtre sur les valeurs du temps de relaxation des systèmes, 869/872 trajectoires sont conservées.

### 3.7.8 Distribution des $R_{||,ij}$

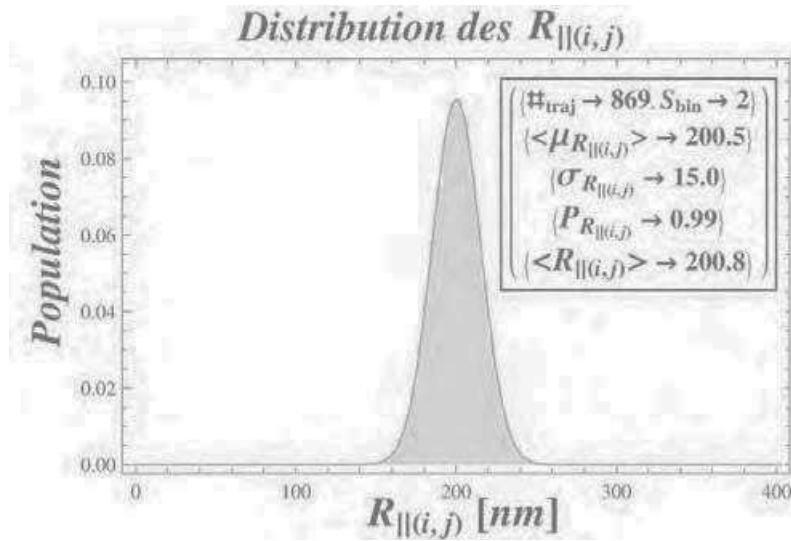


Fig. 3.38 – Distribution des amplitudes du mouvement corrigées par l'effet du temps de corrélation  $R_{||,ij}$ , ( $L_{ADN} = 1201$  bp,  $R_{par} = 150$  nm)

Un ajustement gaussien appliqué à l'ensemble de la distribution de probabilité des  $R_{||,ij}$ , Fig. 3.38, nous permet d'extraire les grandeurs suivantes :

- $\mu_{\textcircled{5}} \Rightarrow \langle \mu_{R_{||,ij}} \rangle_{\textcircled{1}}$
- $\sigma_{\textcircled{5}} \Rightarrow \sigma_{R_{||,ij}} \textcircled{1}$

De plus, nous calculons la moyenne mathématique de l'ensemble des amplitudes

- $\langle \textcircled{1} \rangle \Rightarrow \langle R_{||,ij} \rangle_{\textcircled{1}}$

La concordance entre les valeurs  $\langle \mu_{R_{||,ij}} \rangle_{\textcircled{1}}$  et de  $\langle R_{||,ij} \rangle_{\textcircled{1}}$  est bonne  $\Rightarrow \Delta_{measure\textcircled{1}} < 0.5$  nm, soit un écart relatif inférieure à 0.2%.

### 3.7.9 Filtre sur les $R_{||,ij}$

Lors de certains expériences nécessitant un temps d'acquisition plus important, nous avons pu observer une augmentation de la valeur  $\Rightarrow \Delta_{measure\textcircled{1}} \sim 1$  à 2 nm.

En examinant en détail les trajectoires sélectionnées contenues dans notre distribution finale, nous avons pu mettre en lumière un événement perturbateur, Fig. 3.39.

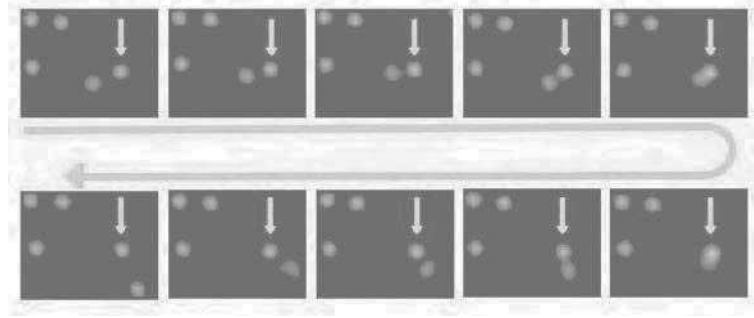


Fig. 3.39 – Illustration en séquence d'image, de perturbation d'une particule sur une autre. ( $L_{ADN} = 1201 \text{ bp}$ ,  $R_{par} = 150 \text{ nm}$ )

En effet, la probabilité de *décrochage* de la surface de nos systèmes ADN/particule n'est pas nulle. Plus la durée d'acquisition augmente plus la probabilité d'observer cet événement augmente.

De plus, ces complexe décrochés soumis au mouvement brownien ne sont plus confinés. Il arrive qu'ils viennent interférer avec d'autres complexes ADN/particule, comme présentés sur la séquence image-par-image ci-dessus. Ces systèmes décrochés viennent "polluer" le comportement des complexes ADN/particule valides, Fig. 3.40

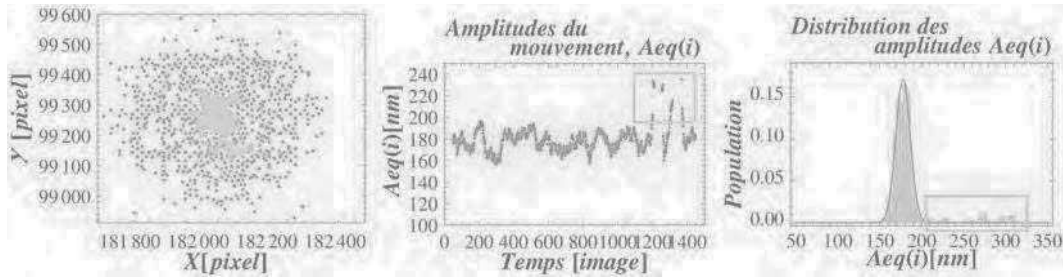


Fig. 3.40 – Implication sur les grandeurs mesurées de cet événement perturbatif d'une particule sur une autre. ( $L_{ADN} = 1201 \text{ bp}$ ,  $R_{par} = 150 \text{ nm}$ )

Cette fois-ci, le filtre ne conduit pas à rejeter l'ensemble d'une trajectoire mais seulement les points de cette dernière qui ne remplissent pas les critères de sélectivité. Seuls, les points,  $R_{||,ij}$ , des trajectoires remplissant le critère de sélectivité suivant sont conservés :

$$\mu_{\textcircled{5}} - 2.5 \times \sigma_{\textcircled{5}} \leq R_{||,ij} \leq \mu_{\textcircled{5}} + 2.5 \times \sigma_{\textcircled{5}} \quad \text{Filtre : } R_{||,ij} \textcircled{1} \quad (3.36)$$

Un ajustement gaussien appliqué à l'ensemble de la distribution de probabilité des  $R_{||,ij}$ , Fig. 3.41, nous permet d'extraire les grandeurs suivantes :

$$\begin{aligned} - \mu_{\textcircled{6}} &\Rightarrow \langle \mu_{R_{||,ij}} \rangle_{\textcircled{2}} \\ - \sigma_{\textcircled{6}} &\Rightarrow \sigma_{R_{||,ij}} \textcircled{2} \end{aligned}$$

De plus, nous calculons la moyenne mathématique de l'ensemble des amplitudes

$$- \langle \textcircled{2} \rangle \Rightarrow \langle R_{||,ij} \rangle_{\textcircled{2}}$$



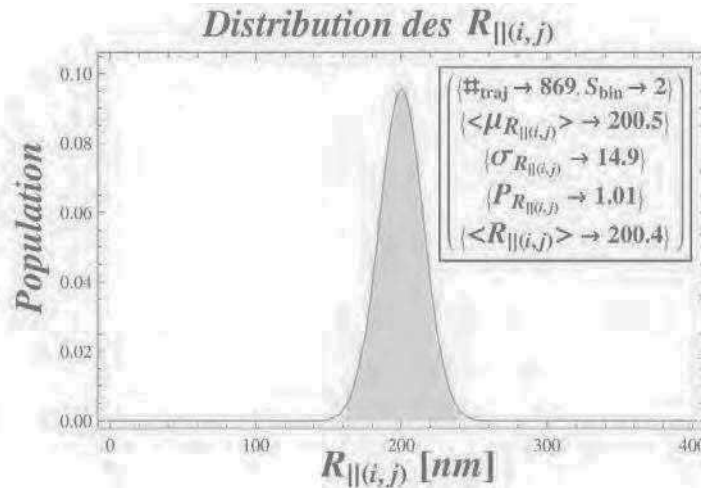


Fig. 3.41 – Distribution des amplitudes du mouvement  $R_{||,i,j}$ , ( $L_{ADN}=1201$  bp,  $R_{par} = 150$  nm)

Les valeurs  $\langle \mu_{R_{||,i,j}} \rangle_{\textcircled{2}}$  et de  $\sigma_{R_{||,i,j}} \textcircled{2}$  concordent à  $\Delta_{\text{mesure}} \textcircled{2} = 0.1$  nm près. L'écart était donc bien dû à ce type d'événement. Suite à l'application de ce filtre sur les valeurs de  $R_{||,i,j}$ , 869 trajectoires sont conservées 1 194 340/1 215 529 points de mesures, soit 21 188 points rejetés.

La distribution en 3 dimensions des positions  $X_{corr,i,j}$  et  $Y_{corr,i,j}$  des trajectoires conservées au cours de ces différentes étapes de traitement, montre une très grande homogénéité de nos mesures, Fig. 3.42.

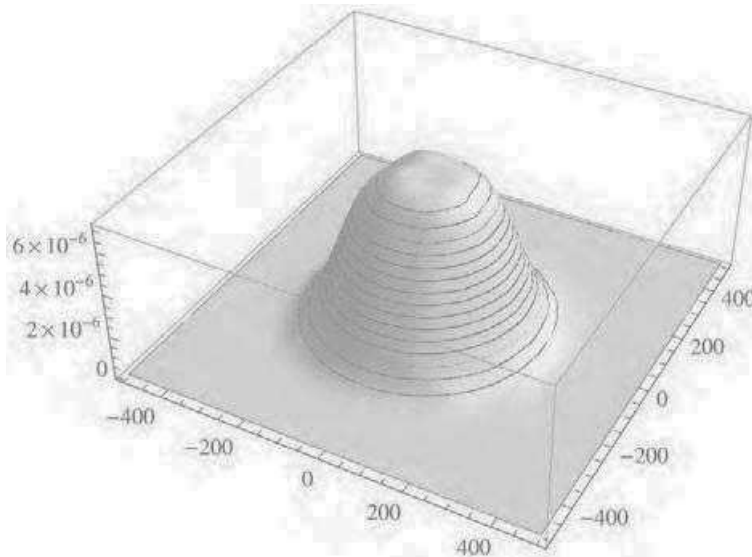


Fig. 3.42 – Distribution de probabilité en 3 dimensions des positions  $X_{corr,i,j}$  et  $Y_{corr,i,j}$  de l'ensemble des trajectoires conservées ( $L_{ADN} = 1201$  bp,  $R_{par} = 150$  nm)

La valeur de l'amplitude du mouvement de nos systèmes ADN/particule avec laquelle nous allons continuer l'analyse des résultats, sont les valeurs  $\langle \mu_{R_{||,i,j}} \rangle_{\textcircled{1}}$  corrigées du temps de corrélation présenté en 3.7.8. Si  $\Delta_{\text{mesure}} \textcircled{1} > 1$  nm, nous appliquons ce dernier filtre. Notre grandeur de travail devient alors  $\langle \mu_{R_{||,i,j}} \rangle_{\textcircled{2}}$ , présentée en 3.7.9

### 3.8 MESURES TPM

L'amplitude du mouvement de nos complexes ADN/particule mesurée grâce au TPM-on-a-chip, nous permet de caractériser certains paramètres physiques de la molécule d'ADN utilisée. Notamment, la longueur bout-à-bout est directement proportionnelle à l'amplitude du mouvement qui quantifie la longueur apparente de la molécule observable par TPM.

#### 3.8.1 Courbe de calibration

La courbe de calibration suivante, met en lumière cette dépendance. Pour réaliser cette dernière, différentes tailles d'ADN sont utilisés. Les paramètres de fonctionnalisation des surfaces, les protéines et anti-corps ainsi que les nanoparticules utilisés sont, bien sur, les mêmes. De plus, les conditions liées au solvant sont identiques.

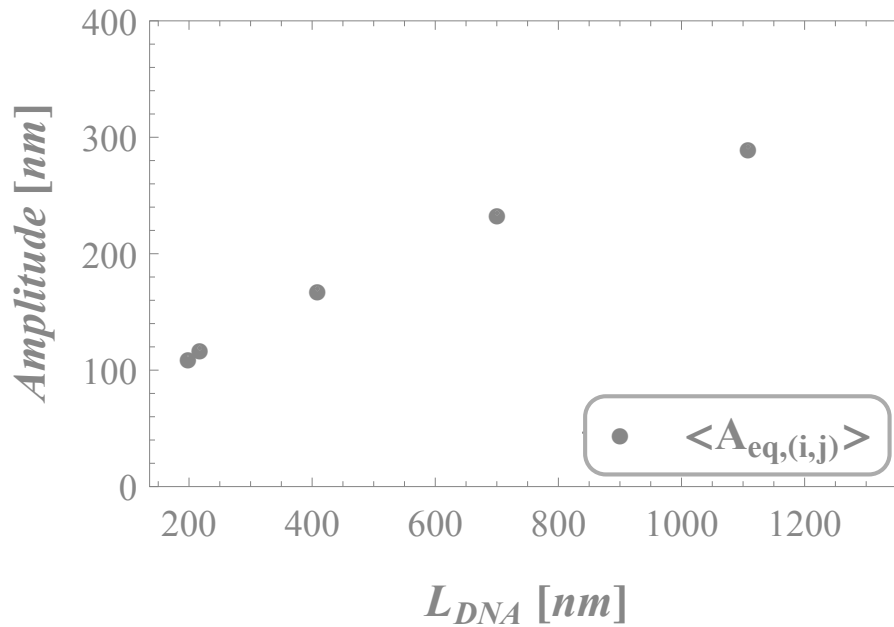


Fig. 3.43 – Courbe de calibration sur les amplitudes brutes,  $\langle \mu_{A_{eq(i,j)}} \rangle$ . ( $L_{ADN} = 583, 639, 1201, 2060, 3259$  bp,  $R_{par} = 150$  nm). Les barres d'erreur sont contenues dans les symboles.

Cette courbe, Fig. 3.43, illustre la dépendance de l'amplitude du mouvement  $\langle \mu_{A_{eq(i,j)}} \rangle$ , que nous noterons dans la suite  $R_{TPM}$  dans un souci de simplicité, en fonction de la longueur de l'ADN.

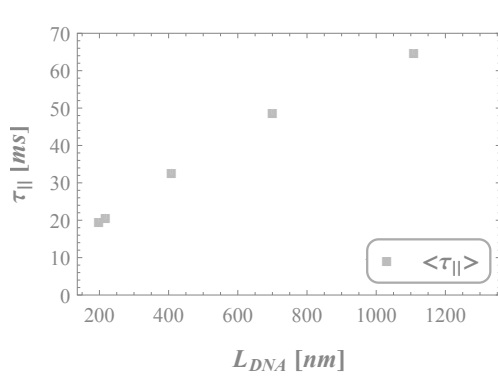


Fig. 3.44 – Courbe de calibration du temps de relaxation des complexes ADN/particule  $\langle \tau_{||} \rangle$ . ( $L_{ADN} = 583, 639, 1201, 2060, 3259$  bp,  $R_{par} = 150$  nm).

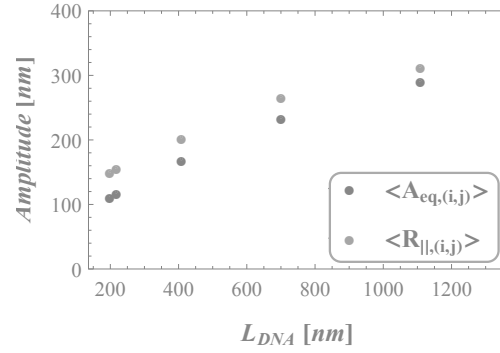


Fig. 3.45 – Courbe de calibration sur les amplitudes brutes,  $\langle \mu_{A_{eq,i,j}} \rangle$  ainsi que les amplitudes corrigées de l'effet lié au temps de corrélation,  $\langle \mu_{R_{||,i,j}} \rangle$ . ( $L_{ADN} = 583, 639, 1201, 2060, 3259$  bp,  $R_{par} = 150$  nm).

Nous pouvons aussi représenter la dépendance du temps de relaxation  $\tau_{||}$ , des complexes ADN/particules en fonction de la longueur de l'ADN, Fig. 3.44. Enfin, nous représentons la variation des amplitudes corrigées de l'effet de moyennage vidéo,  $\langle \mu_{R_{||,i,j}} \rangle$  que nous noterons dans la suite  $R_{Exp||}$ , en fonction de la longueur de l'ADN, Fig. 3.45.

Ces courbes indiquent également le minimum de variation de l'amplitude du mouvement de nos complexes ADN/particule détectable par TPM.

### 3.8.2 Extraction de la contribution de l'ADN seul

Les mesures TPM, nous permettent d'extraire l'amplitude du mouvement des complexes ADN/particule observés en 2 dimensions,  $\langle R_{Exp||} \rangle_{2D}$ . Il est nécessaire, en premier lieu, de revenir à la contribution 3D de nos mesures afin d'extraire la longueur bout-à-bout de la molécule d'ADN. Cette caractéristique physique peut être extraite via deux méthodes.

#### Méthode minimale

Cette expression repose sur le fait que nous pouvons considérer la particule et l'ADN comme deux maillons statistiquement indépendants. Pour extraire la distance bout-à-bout de la molécule d'ADN nous appliquons alors, la simple correction suivante :

$$\langle R_{Exp||}^2 \rangle_{3D} = \frac{3}{2} \langle R_{Exp||}^2 \rangle_{2D} \quad (3.37)$$

Ensuite, la contribution de la particule à l'amplitude du mouvement est soustraite, Eq. 3.38. La dépendance de  $R_{ADN}$  en fonction de  $L_{ADN}$  est représentée en Fig. 3.46.

$$\begin{aligned} \langle R_{ADN}^2 \rangle &= \langle R_{Exp||}^2 \rangle_{3D} - R_{par}^2 \\ \Rightarrow \langle R_{ADN}^2 \rangle &= \frac{3}{2} \langle R_{Exp||}^2 \rangle_{2D} - R_{par}^2 \end{aligned} \quad (3.38)$$

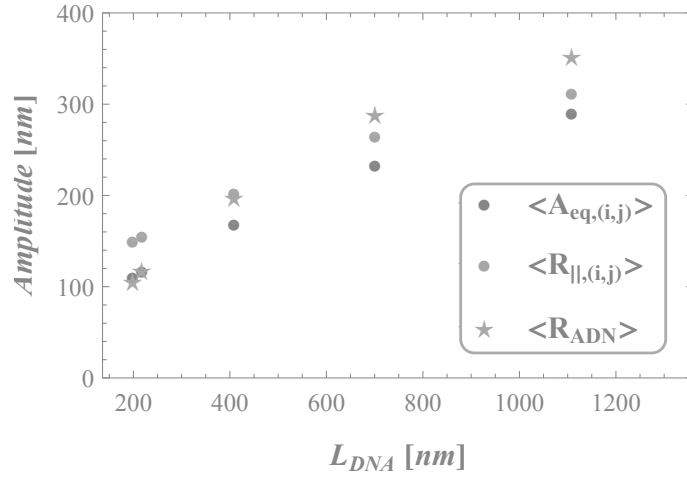


Fig. 3.46 – Courbe de calibration de la longueur bout-à-bout de la molécule d'ADN,  $\langle \mu_{R_{DNA}} \rangle$ . ( $L_{ADN}=583,639,1201,2060,3259$  bp,  $R_{par} = 150$  nm).

#### Méthode de Segall et al.

Cette méthode est proposée dans le papier [SNPo6]. Elle doit être appliquée lorsque le nombre d'exclusions est  $N_R \gg 1$ . En supposant que la chaîne de polymère suit un comportement de régime gaussien dans Eq. (10c) de Segall 2006, l'extraction de la distance bout-à-bout de l'ADN se calcule par la résolution de l'équation suivante :

$$\langle R_{Exp||}^2 \rangle_{2D} = \frac{\langle R_{ADN}^2 \rangle}{2} + \sqrt{\frac{8}{3\pi}} \cdot \frac{R_{par} \cdot \sqrt{\langle R_{ADN}^2 \rangle}}{\text{erf}\left[\frac{R_{par} \cdot \sqrt{6}}{\sqrt{\langle R_{ADN}^2 \rangle}}\right]} \quad (3.39)$$

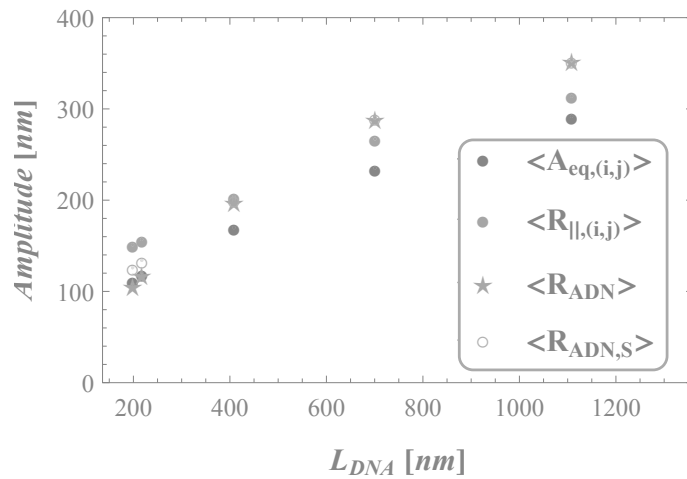


Fig. 3.47 – Courbe de calibration de la longueur bout-à-bout de la molécule d'ADN,  $\langle \mu_{R_{DNA}} \rangle$ . ( $L_{ADN} = 583,639,1201,2060,3259$  bp,  $R_{par} = 150$  nm).

Comme illustré sur le graphique, Fig. 3.47, il convient d'utiliser la méthode de Segall pour extraire la distance bout-à-bout des ADNs de petite taille.

Cette équation prend en compte le phénomène de volume exclus dû à la présence de la nanoparticule, notamment lorsque cette dernière est proche de la surface solide, ce qui contraint le domaine de conformations accessible au polymère et induit, sur ce dernier, une force verticale effective d'étirement d'origine entropique.

### 3.8.3 Extraction de la longueur de persistance

Une fois la distance bout-à-bout de la molécule d'ADN calculée, nous pouvons utiliser les modèles de physique statistique afin d'extraire la longueur de persistance.

Pour rappel, en utilisant les modèles **FJC** et **WLC** l'extraction de  $L_p$  se fait par la résolution des équations suivantes :

$$\langle R_{ADN}^2 \rangle_{3D}^{FJC} = 2LL_{p,FJC} \quad (3.40)$$

$$\langle R_{ADN}^2 \rangle_{3D}^{WLC} = 2LL_{p,WLC} - L_{p,WLC}^2 \left( 1 - \exp \left[ -\frac{L}{L_{p,WLC}} \right] \right) \quad (3.41)$$

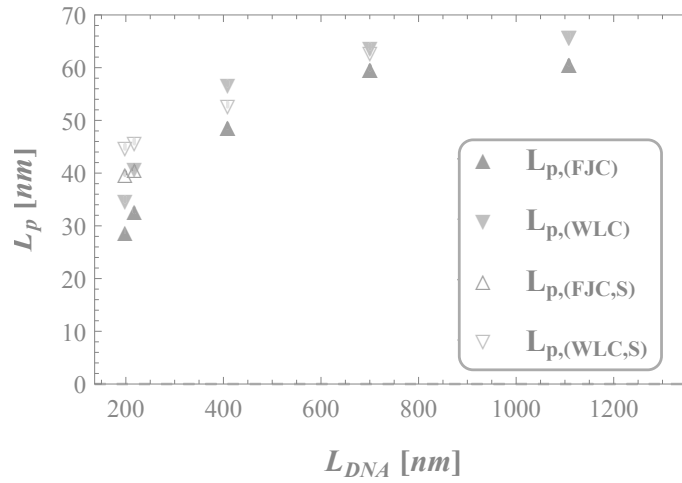


Fig. 3.48 – Courbe de calibration de la longueur de persistance de la molécule d'ADN,  $L_{p,(FJC)}$  et  $L_{p,(WLC)}$  en utilisant respectivement les modèles **FJC** et **WLC**, les valeurs de la distance bout-à-bout sont extraies soit par la méthode minimale soit par la méthode Segall (symboles vides) (ADN  $L_{ADN}=583,639,1201,2060,3259$  bp,  $R_{par} = 150$  nm).

Nos mesures TPM nous permettent de remonter jusqu'aux propriétés physiques de la molécule d'ADN. Notamment, l'estimation de sa caractéristique intrinsèque qu'est la longueur de persistance, Fig. 3.48.

## CONCLUSION DU CHAPITRE

En plus de son importance biologique indéniable, l'ADN est un objet d'intérêt croissant. En effet, en tant que modèle expérimental, il permet de rendre compte du comportement statistique d'une chaîne de polymères et d'affiner les modèles théoriques. Les techniques de suivi de molécule unique ont contribué aux grandes avancées dans cette voie. Ces études permettent d'explorer et d'approfondir le comportement dynamique de l'ADN. La technique de suivi de molécule unique que nous utilisons est le TPM. La force appliquée à la molécule d'ADN via cette méthode est extrêmement faible afin de ne pas limiter les degrés de liberté conformationnel du système. Ainsi, cela nous permet de sonder la réponse élastique de l'ADN

Afin d'obtenir des résultats ayant un poids statistique suffisamment pertinent, une bio-puce, nommée TPM-on-a-chip, fut mise en place et la technique associée fut brevetée par l'IPBS. Cette bio-puce nous offre la possibilité de suivre simultanément le mouvement brownien de plusieurs centaines de molécules d'ADN individuelles. Cette dernière est associée à un protocole d'analyse de données rigoureux et créé spécifiquement pour ces expériences. Tout ceci nous permet d'éliminer les traces erratiques afin d'accumuler un grand nombre de trajectoires physiquement cohérentes.

Ainsi, nous pouvons extraire les propriétés physiques des complexes ADN/particule tels que leur longueur apparente ou encore le temps de relaxation du système, qui sont tous deux directement dépendants de la longueur de la molécule d'ADN utilisée. Via l'utilisation des modèles de physique statistique classique, nous pouvons remonter à la longueur apparente de la molécule d'ADN ainsi qu'à sa longueur de persistance. En fonction du régime de polymère utilisé, autrement dit en fonction de la taille de la molécule d'ADN, la méthode d'extraction et les modèles employés sont à adapter. En effet, nous avons pu percevoir les différences induites lors de la soustraction de la contribution de la particule au mouvement brownien, ou encore les différences entre les modèles **FJC** et **WLC** lors de l'extraction de  $L_p$ . Nous aborderons de nouveau ces questions dans la suite.

Nous pouvons prospecter l'impact de contraintes globales, ou locales sur les changements conformationnels de la molécule. De plus, nous pouvons explorer l'effet de ces changements sur la propriété physique principale de la molécule d'ADN qu'est la longueur de persistance. Pour résumer, cette technique nous permet donc de remonter aux propriétés intrinsèques de la molécule d'ADN.

## SOMMAIRE

INTRODUCTION . . . . .	107
4.1 SIMULATION MONTE-CARLO DYNAMIQUE (MC-D) . . . . .	109
4.1.1 Simulation du polymère . . . . .	109
Longueur du polymère . . . . .	109
Longueur de persistance . . . . .	110
Temps de calcul . . . . .	111
4.1.2 Simulation de la nanoparticule . . . . .	111
4.1.3 Simulation d'une chaîne réelle et volume exclu . . . . .	111
Sphères dures . . . . .	111
Potentiel de Lennard-Jones . . . . .	112
4.1.4 Simulation des énergies mises en jeu dans le système . . . . .	113
Module de courbure et d'étirement . . . . .	113
Énergie de courbure et d'étirement . . . . .	114
4.1.5 Simulation des contraintes expérimentales . . . . .	114
Murs durs . . . . .	114
Représentation de la simulation . . . . .	115
4.1.6 Évaluation du mouvement lors d'un <i>pas</i> Monte-carlo, <i>pMC</i> . . . . .	115
Phase d'initialisation . . . . .	115
Définition du mouvement par tirage aléatoire . . . . .	115
<i>Pas</i> spatial . . . . .	116
Principe d'ergodicité . . . . .	116
Lien avec le mouvement brownien . . . . .	116
Durée d'un <i>pas</i> Monte-Carlo . . . . .	117
Acceptation ou refus des configurations . . . . .	117
Résumé d'un <i>pas</i> Monte-Carlo . . . . .	118
4.1.7 Évaluation du mouvement lors d'un balayage Monte-carlo, <i>bMC</i> . . . . .	118
Durée d'une <i>boucle</i> Monte-Carlo . . . . .	119
Durée d'une simulation Monte-Carlo . . . . .	119
Fréquence d'observation . . . . .	120
4.1.8 Exploitation des données . . . . .	120
Extraction des données . . . . .	120
Courbe de calibration . . . . .	120
Influence de la discrétisation . . . . .	121
Influence du volume exclu . . . . .	123
4.1.9 Temps de corrélation . . . . .	124

4.2	SIMULATION PAR ÉCHANTILLONNAGE EXACT (EE)	125
4.2.1	Simulation des contraintes expérimentales	126
	Sphères dures et volume exclu	126
	Murs durs	126
4.2.2	Évaluation des conformations du système	126
	Monomère et segment initial	126
	Autres monomères	127
	Nanoparticule	127
4.2.3	Exploitation des données	127
	Extraction des données	127
	Courbe de calibration	128
	Acceptation ou refus des configurations	130
	CONCLUSION	131



## INTRODUCTION

LES outils numériques de simulation sont aujourd’hui des outils incontournables pour la construction et le développement des connaissances scientifiques. Leur intérêt est multiple et se situe sur plusieurs plans. En effet, les simulations numériques permettent d’explorer et d’étudier des domaines où la théorie seule n’apporte pas de réponse. Ceci est souvent dû au nombre de simplifications qui justement sont liées à l’élaboration des modèles. La simulation permet aussi de modéliser les systèmes expérimentaux. L’étape de rationalisation des paramètres clés qui y est associée est donc extrêmement importante. Enfin, la simulation numérique donne la possibilité de tester les approximations des modèles théoriques. Le fait de corrélérer et confronter de façon permanente les outils et les résultats expérimentaux, théoriques et numériques permet une amélioration continue de ces derniers. En effet, les allers-retours incessants entre ces trois approches permettent d’améliorer la compréhension et la production scientifique.

L’expression et la représentation fidèles des phénomènes observés induisent rapidement des difficultés, notamment dues aux calculs non explicites et à l’absence d’informations sur certains paramètres. En effet, certaines données ne sont pas accessibles. Par exemple, l’estimation directe de l’énergie libre par des méthodes statistiques est impossible si nous ne connaissons pas explicitement la densité de probabilité, alors que l’estimation de la différence d’énergies libres est accessible. Certaines techniques de simulation permettent plus rapidement d’approcher numériquement ce dernier résultat. Durant ma thèse, nous avons utilisé une approche numérique basée sur la méthode de Monte-Carlo (MC).

La méthode de simulation Monte-Carlo Metropolis fut introduite pour la première fois en tant qu’outil majeur d’investigation en 1953 [MRR<sup>+</sup>53]. Initialement utilisée dans le domaine de la physique de la matière condensée pour l’étude de la diffusion des neutrons dans un matériau fissile, elle s’est étendue, depuis, à de nombreux autres ensembles statistiques. Basée sur l’utilisation de nombres aléatoires, d’où lui provient ce nom<sup>1</sup>, et d’expériences répétées, cette méthode permet l’évaluation de moyennes de grandeurs physiques dans un ensemble statistique.

Nous avons donc utilisé une simulation numérique afin de modéliser les changements conformationnels de la molécule d’ADN en imitant notre géométrie expérimentale ainsi que les contraintes associées. La molécule d’ADN est simulée à l’échelle mésoscopique, l’échelle intermédiaire entre le macroscopique et le microscopique. Autrement dit, le détail atomique de la molécule d’ADN ainsi que des nucléotides la constituant ne sont pas pris en compte. De plus, nous nous intéressons aussi à la dynamique de notre système. En effet, comme nous l’avons vu lors de l’application de la correction de l’effet de moyennage vidéo, le temps de relaxation de nos complexes ADN/particule joue un rôle important sur les propriétés physiques que nous pouvons extraire expérimentalement. Pour rendre compte de cela, nous exploitons un modèle de simulation Monte-Carlo dynamique. Dans ce chapitre nous allons décrire et préciser la construction et les paramètres clés de ce modèle.

---

1. Méthode de simulation désignée *Monte-Carlo*, en référence aux jeux de hasard des casinos



## 4.1 SIMULATION MONTE-CARLO DYNAMIQUE (MC-D)

Notre modèle de simulation est un outil capable de tester différents paramètres intervenant sur les changements conformationnels de la molécule d'ADN. Comme nous l'avons vu précédemment, la dynamique de certains événements ne nous est pas accessible par nos mesures expérimentales par TPM. Notre modèle de simulation **Monte-Carlo Dynamique**, MC-D, est construit afin d'être un outil particulier capable de sonder la dynamique de certains événements intervenant sur nos systèmes ADN/particule [MTB<sup>+</sup><sub>10</sub>].

### 4.1.1 Simulation du polymère

#### Longueur du polymère

Le polymère d'une longueur  $L$  est défini tel une suite de  $N$  **monomères** de rayon  $a$  à laquelle est accrochée une nanoparticule de rayon  $R_{par}$ . Ainsi, chaque **monomère** représente un ensemble de plusieurs paires de bases. Un monomère ne caractérise donc pas la nature des bases azotés le constituant, ni les détails atomiques et moléculaires de l'ADN.

En pratique, l'indexation des  $N$  monomères dans le programme codé en *language C*, est réalisée telle que  $i \in [0 : N - 1]$  ainsi, il en découle que la nanoparticule est indexée en  $i = N - 1$ , comme représenté sur le schéma Fig. 4.1.

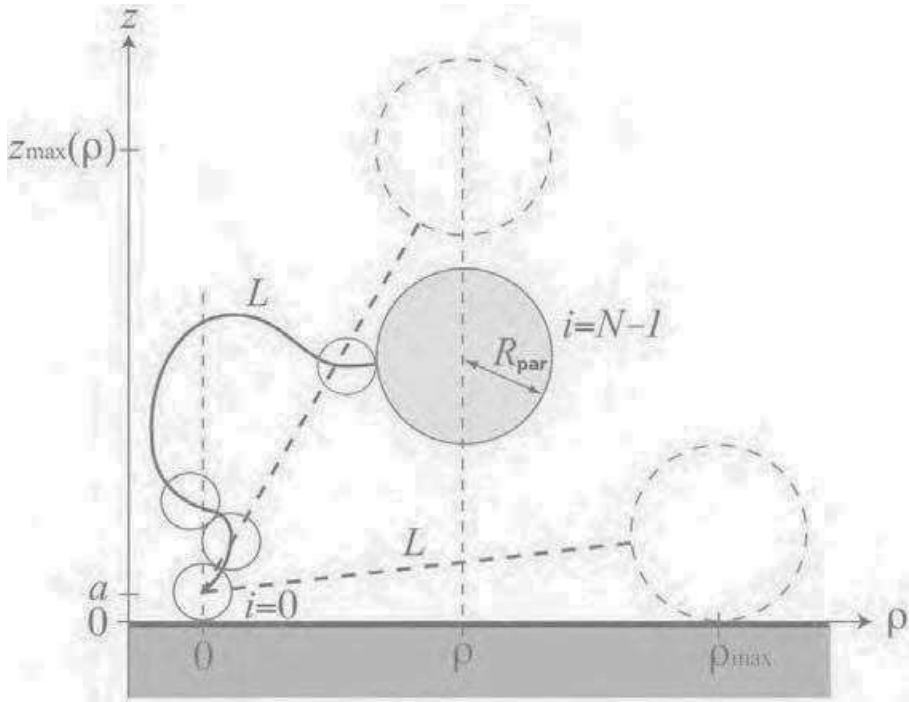


Fig. 4.1 – Schéma de la conformation et des paramètres de notre simulation. Le polymère, de longueur de contour  $L$ , est modélisée par une succession de  $N$  monomères de rayon  $a$ , au-bout duquel est accrochée la nanoparticule modélisée par une sphère de rayon  $R_{par}$ . La projection sur le plan  $(X, Y)$  de la distance bout-à-bout du complexe ADN/particule est notée  $\rho$ , et  $\rho_{max}$  est l'extension maximale du complexe.

La molécule d'ADN est modélisée telle un "collier de perle", ce type de simulation est aussi nommé simulation en *gros grain* (en : *coarse-grained*). La longueur du polymère  $L$  définie dans le programme correspond donc à :

$$L = 2 \times a \cdot (N - 1) \quad (4.1)$$

La fixation de l'ADN sur la surface solide est modélisée via un point d'ancrage fixe. Le centre du premier monomère  $i = 0$ , est donc considéré comme un point, fixé à l'origine dans le plan (X,Y). Ce dernier est placé selon l'axe Z tel que  $z(i = 0) = a$ . La demi-sphère inférieure du monomère  $i = 0$  n'est pas prise en compte lors de la mesure de la distance bout-à-bout du système.

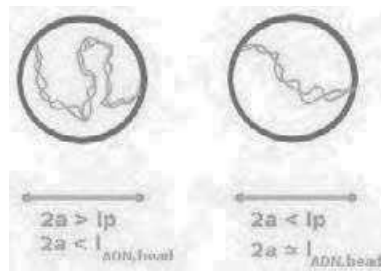
De nombreuses simulations furent réalisées, en amont de ma thèse, afin de tester l'influence d'un point d'ancrage rigide ou d'un segment FJC sur la modélisation du mouvement. Aucune différence significative n'a été observée, la modélisation la plus simple, correspondant à un point d'ancrage FJC, a donc été conservée. De plus, pour l'ADN réel, à l'heure actuelle, nous savons pas exactement quelle est la nature de l'ancrage de ce dernier.

Nous avons une certaine liberté sur le choix du paramètre  $N$ , mais il est nécessaire de respecter quelques règles afin que la simulation du polymère soit physiquement cohérente. En effet, une molécule d'ADN de 2060 bp, dont le comportement se rapproche du modèle gaussien, ne peut certainement pas être modélisée par seulement  $N = 2$  monomères ! La nature même du polymère en serait altérée. Nous allons détailler les contraintes et limites liées à la discrétisation de la molécule d'ADN.

### Longueur de persistance

Par défaut la longueur de persistance de notre modèle est prise égale à la valeur de 50 nm [Hag88], soit  $L_p = 147$  bp, valeur associée à la longueur de persistance d'un ADN double brin dans des conditions physiologiques de température<sup>2</sup> et de salinité<sup>3</sup>. Autrement dit,  $L_p = L_{p_{Ref}}/0.34 = 147$  bp.

Le diamètre des monomères,  $2a$ , doit rester strictement inférieur à la longueur de persistance  $L_p$ , sinon l'extrapolation de la longueur de l'ADN à l'intérieur du monomère  $l_{ADN,bead}$  est biaisé. En effet, si  $2a > L_p$  à l'échelle d'un monomère, l'ADN est en régime flexible. Ceci implique que l'ADN peut se recourber, s'enrouler ou se mettre en pelote dans le monomère, comme schématisé ci-dessous Fig. 4.2.



Il est donc nécessaire de respecter la limite suivante :

$$2a \ll L_p \quad (4.2)$$

$$\Rightarrow \frac{L}{(N - 1)} \ll L_p \quad (4.3)$$

$$\Rightarrow (N - 1) \gg \frac{L}{L_p} \quad (4.4)$$

Fig. 4.2 – Comparaison de la conformation de l'ADN dans un monomère en fonction de son rayon  $a$

2.  $T_{physio} = 37.5$  °C (310.5 K)

3.  $C[NaCl]_{physio} = 0.9$  %  $\Rightarrow 155$  mM

### Temps de calcul

La limite supérieure pour la valeur de  $N$  dépend du temps de calcul que nous acceptons pour nos simulations. En effet, plus la valeur de  $N$  est grande, donc plus le nombre de paire de bases par monomère diminue, plus le temps de calcul des conformations accessibles à la chaîne augmente.

De façon générale [MTB<sup>+</sup>10], nous nous donnons les bornes suivantes pour la valeur du diamètre d'un monomère :

$$8 \text{ bp} < 2a < 42 \text{ bp} \quad (4.5)$$

$$\Rightarrow 2.8 \text{ nm} < 2a < 14 \text{ nm} \quad (4.6)$$

Ceci permet de respecter les propriétés physiques de l'ADN étudié et de ne pas lancer des simulations trop coûteuses en temps de calcul.

#### 4.1.2 Simulation de la nanoparticule

La nanoparticule est fixée au dernier maillon constituant la chaîne du polymère. Son rayon,  $R_{par}$ , est défini en fonction de la taille des nanoparticules utilisées durant nos expériences TPM. Cette dernière est considérée comme un objet dur, un objet dont le volume est impénétrable par tout autre objet du système.

#### 4.1.3 Simulation d'une chaîne réelle et volume exclu

Nous cherchons à rendre compte des interactions répulsives existant entre les différents objets de notre système. Nous appliquons des conditions à courte et longue distance, afin d'interdire l'interpénétration de la chaîne avec elle-même, et ainsi rendre compte des effets dus aux volumes exclus dans un bon solvant. De plus, nous considérons la particule comme une sphère dure. Pour tout cela nous appliquons deux types de répulsions.

- La répulsion due au volume d'exclusion existant entre voisins immédiats le long de la chaîne, monomères indexés  $i - 1, i$  et  $i + 1$ . La prise en compte de ces conditions d'exclusions est réalisée en considérant chaque monomère ainsi que la particule, comme un d'objet dur.
- La répulsion à longue distance d'un monomère avec la chaîne ou avec la particule, est introduite via le potentiel de Lennard-Jones. Cette répulsion est appliquée lorsque les objets deviennent spatialement proches, objets indexés  $i$  et  $j$  pour tout  $j \notin \{i - 1, i, i + 1\}$

L'influence de ces répulsions est évaluée entre la position antérieure au déplacement et la position du système si le mouvement avait été accepté. Le détail de la prise en compte du volume d'exclusion autour de chaque monomère est présenté ci-dessous.

### Sphères dures

Chaque monomère est considéré comme une sphère dure de rayon  $a$  auquel est associé un volume d'exclusion,  $V_{exc} = \frac{4\pi}{3}a^3$ . Ce critère de **sphère dure** est mesuré en objets proches voisins. Il est pris en compte lors du calcul et de l'acceptation du

mouvement de chacun des monomères. Il en résulte, qu'à chaque instant  $t$ , aucun monomère ne peut pénétrer le volume occupé par un autre monomère.

Ce paramètre est utilisé lors de la mesure du mouvement entre voisins immédiats, mais aussi entre voisins spatiaux comme nous allons le voir.

### Potentiel de Lennard-Jones

Le modèle *gros grain* que nous utilisons, s'affranchit des détails à l'échelle atomique. Ceci permet de réduire la complexité du système. Afin de rendre compte des degrés de liberté et des limitations qui leur sont associés, les potentiels d'interaction effectifs intervenant entre les unités du modèle sont donc ajoutés.

Les interactions entre les monomères sont modélisées par un potentiel de **Lennard-Jones** qui prend généralement en compte les effets attractifs et répulsifs à courte distance entre deux objets [Jon24].

$$U_{LJ} = \sum_{i < j} \left[ \frac{b}{(|\mathbf{r}_i - \mathbf{r}_j| - 2a)^{12}} - 2 \left( \frac{b}{(|\mathbf{r}_i - \mathbf{r}_j| - 2a)^6} \right) \right] \quad (4.7)$$

Lorsque les deux corps sont trop proches, la contribution répulsive du potentiel est dominante. Tandis que lorsque les deux corps s'éloignent l'interaction attractive existant entre eux prend le dessus. La distance d'équilibre,  $r_0$ , entre ces deux objets, où les forces attractives et répulsives se compensent, correspond au puits de potentiel énergétique, Fig. 4.3 .

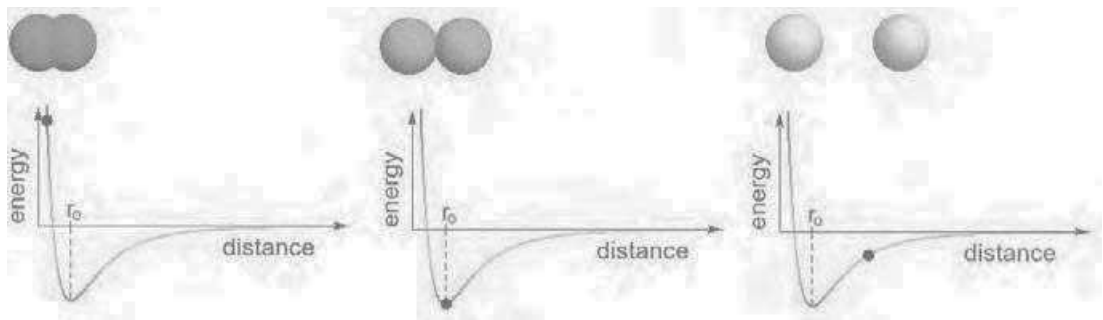


Fig. 4.3 – Schéma du potentiel de Lennard-Jones à différentes distances d'interactions.

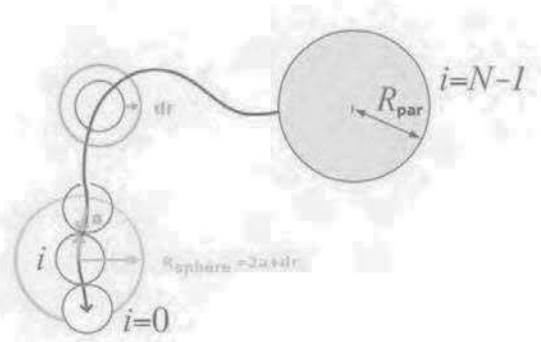
Ce potentiel va nous permettre de traduire les **interactions répulsives** lorsque deux objets deviennent proches spatialement. Pour cela nous tronquons ici ce potentiel dans la simulation afin de ne prendre en compte que sa partie répulsive pour  $r < r_0$ . Le potentiel de Lennard-Jones suffit à rendre compte efficacement des interactions de notre modèle sans être trop coûteux en temps de calcul.

La zone d'effet du potentiel dépend du volume occupé par les objets de notre système. De plus, il est nécessaire de tester l'influence de ce potentiel sur notre objet, avant et après que ce dernier ait subi un déplacement. Le déplacement d'un monomère lors d'une itération de la simulation est  $dr$ . Les *zones d'interaction* liées à l'application du potentiel de Lennard-Jones, sont définies autour de chacun des objets de notre système. L'étendue de ces zones dépend de la nature des objets qui sont en interaction, soit monomère-monomère ou monomère-particule.

Le mouvement pourra être accepté si la *zone d'interaction*, définie ci-dessous, est accessible avant le mouvement. Un test est donc effectué sur les volumes disponibles autour de chaque objet, selon les limites suivantes :

- La *zone d'interaction* définie autour d'un monomère  $i$ , dépend du volume occupé par un monomère augmenté du volume associé à l'amplitude de son éventuel déplacement aléatoire. La taille de cette *zone d'interaction* est une sphère de rayon  $R_{sphere}$ , définie par :

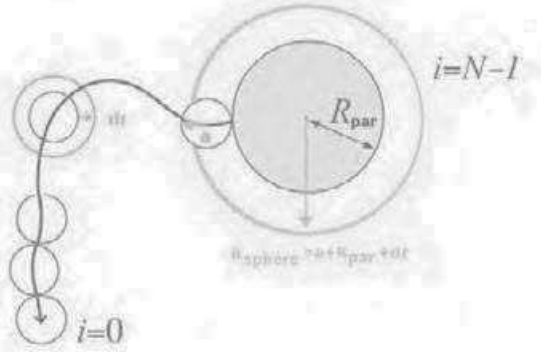
$$R_{sphere} = 2a + dr \quad (4.8)$$



Si un objet  $j \in [0 : N - 1]$  tel que  $j \neq \{i - 1, i, i + 1\}$ , est situé dans ce volume défini autour du monomère  $i$ , alors le potentiel de Lennard-Jones est calculé afin de rendre compte de l'interaction répulsive existant entre ces objets.

- La *zone d'interaction* définie autour de la particule dépend là encore, des mêmes règles. La taille de cette zone est la sphère définie par :

$$R_{sphere} = a + R_{par} + dr \quad (4.9)$$



Si un objet  $i \in [0 : N - 2]$ , est situé dans ce volume défini autour de la nanoparticule  $i = N - 1$ , alors le potentiel de Lennard-Jones est calculé afin de rendre compte de l'interaction répulsive existant entre ces objets.

Ce potentiel d'interaction rend compte des interactions existantes entre deux objets proches spatialement. Si les positions de deux objets de notre système physique sont incluses dans ces *zones d'interaction*, alors le potentiel de Lennard-Jones est calculé.

Cet outil complète la définition des objets de notre système en tant que sphères dures afin de modéliser plus finement les effets d'exclusion.

#### 4.1.4 Simulation des énergies mises en jeu dans le système

##### Module de courbure et d'étirement

Le module de courbure est défini comme en Eq. 2.18, soit :

$$\kappa_c = L_p k_B T$$

Le module d'étirement est défini comme la solution exacte pour un cylindre élastique isotrope ayant pour rayon  $a$ , soit :

$$\gamma = \frac{4\kappa_c}{a^2} \quad (4.10)$$

$$= \frac{4L_p k_B T}{a^2} \quad (4.11)$$

### Énergie de courbure et d'étirement

Les interactions entre les monomères sont régies par les **énergies de courbures** et **énergies d'étirements**. L'expression de ces énergies diffère selon les objets considérés, monomère-monomère ou monomère-particule.

L'énergie interne du système à température constante, s'exprime par :

$$\begin{aligned}
 U = & \underbrace{\sum_{i=0}^{N-2} \left[ \frac{\kappa_c}{2a} (1 - \cos\theta_i) \right]}_{\text{Énergie de courbure}} - \underbrace{\sum_{i=0}^{N-2} \left[ \frac{\gamma}{4a} (|\mathbf{r}_i - \mathbf{r}_{i+1}| - 2a)^2 \right]}_{\text{Énergie d'étirement}} \\
 & \underbrace{\hspace{10em}}_{\text{Monomère-Monomère}} \\
 + & \underbrace{\left[ \frac{\kappa_c}{2a} (1 - \cos\theta_{N-2}) \right]}_{\text{Énergie de courbure}} - \underbrace{\left[ \frac{\gamma}{4a} (|\mathbf{r}_{N-2} - \mathbf{r}_{N-1}| - R_{par} - a)^2 \right]}_{\text{Énergie d'étirement}} \\
 & \underbrace{\hspace{10em}}_{\text{Monomère-Particule}}
 \end{aligned} \quad (4.12)$$

où chaque segment, défini grâce au centre des monomères successifs, forme un angle  $\theta_i$  avec le segment que le précède.

#### 4.1.5 Simulation des contraintes expérimentales

##### Murs durs

Afin de rendre compte de la présence d'une surface solide dans notre dispositif expérimental, nous appliquons des conditions de **murs durs**.

Lors de l'application d'un déplacement d'un objet de notre système, les conditions de déplacements selon l'axe des  $Z$  sont testées. En effet, si l'objet subit un déplacement l'amenant au voisinage immédiat de la surface, ce mouvement est directement refusé.

Pour cela la distance entre la surface et le centre de l'objet doit rester supérieure à  $z_{max}$ , limites définies selon les conditions suivantes :

- pour les monomères :  $z_{max} = a$
- pour la particule :  $z_{max} = R_{par}$



### Représentation de la simulation

Ainsi simulé, notre complexe numérique ADN/particule modélise notre système expérimental, comme l'illustre la Fig. 4.4.

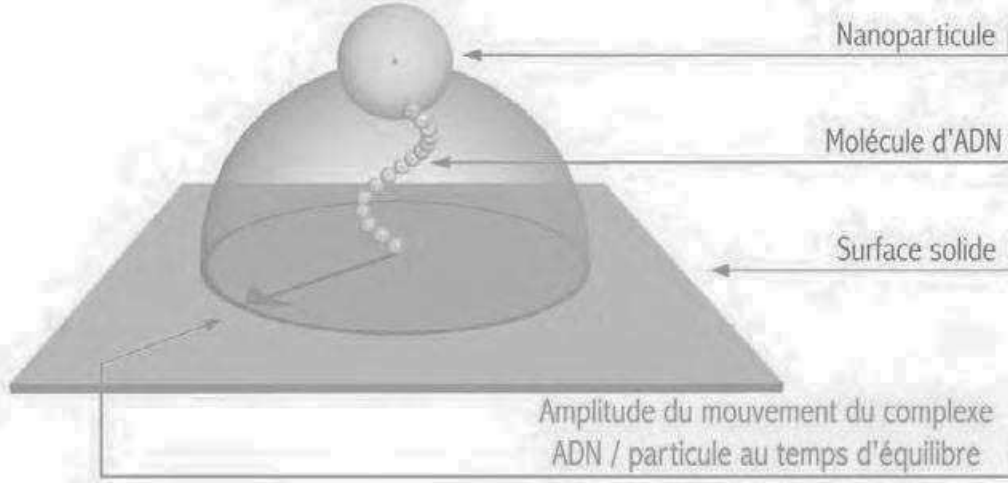


Fig. 4.4 – Schéma du principe d'un complexe numérique ADN/particule modélisant les mesures TPM

#### 4.1.6 Évaluation du mouvement lors d'un pas Monte-carlo, $pMC$

##### Phase d'initialisation

À l'instant initial, le système ADN/particule est en position étirée selon l'axe des  $Z$ , Fig. 4.5. Cette conformation initiale souligne la nécessité d'une durée de simulation adaptée afin de laisser le système relaxer les contraintes conformationnelles et explorer l'espace des états qui lui est accessible.

Chaque monomère ainsi que la particule, sont donc centrés de la façon suivante :

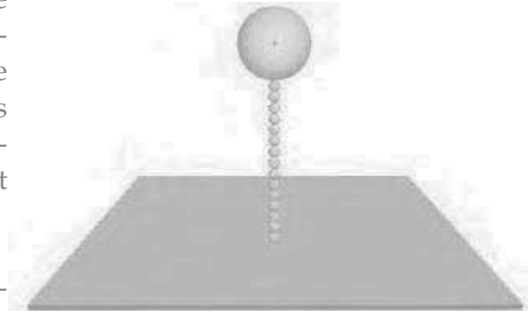


Fig. 4.5 – Conformation initiale

$$\text{Pour } i \in [0 : N - 2] \Rightarrow (x_i, y_i, z_i) = (0, 0, 2a \times i + a) \quad \text{Monomère} \quad (4.13)$$

$$\text{Pour } i = N - 1 \Rightarrow (x_i, y_i, z_i) = (0, 0, 2a \times i + a + R_{par}) \quad \text{Particule} \quad (4.14)$$

##### Définition du mouvement par tirage aléatoire

À chaque pas Monte-Carlo,  $pMC$ , deux processus aléatoires sont effectués. Tout d'abord, un objet parmi les  $N$  objets de notre système, incluant donc monomères et particule, est choisi par tirage aléatoire. Cet objet  $i_{rand}$  subit alors un mouvement construit, là aussi, par tirage aléatoire. Le déplacement est défini selon les trois directions de l'espace, tel que :

$$dr_{rand}^2(i_{rand}) = dx_{rand}^2 + dy_{rand}^2 + dz_{rand}^2 \quad (4.15)$$

### Pas spatial

Le déplacement aléatoire  $dr_{rand}$ , est uniformément distribué dans une sphère de rayon  $R_{dr_{rand}}$ . Ce déplacement statistique possède une valeur moyenne nulle  $\langle dr_{rand} \rangle = 0$ , comme pour la distance bout-à-bout. Son écart quadratique moyen s'exprime selon le rapport des intégrales suivant :

$$\langle dr_{rand}^2 \rangle = \frac{4\pi \int_0^{R_{dr_{rand}}} r^4 dr}{4\pi \int_0^{R_{dr_{rand}}} r^2 dr} \quad (4.16)$$

$$\langle dr_{rand}^2 \rangle = \frac{3}{5} R_{dr_{rand}}^2 \quad (4.17)$$

Dans nos simulations, nous avons fixé le paramètre  $\langle dr_{rand}^2 \rangle$ , [MTB<sup>+</sup><sub>10</sub>], tel que :

$$\Rightarrow \sqrt{\langle dr_{rand}^2 \rangle} = \frac{a}{5} \quad (4.18)$$

Ainsi, cela respecte raisonnablement la contrainte  $\sqrt{\langle dr_{rand}^2 \rangle} \ll a$ , autrement dit la valeur du déplacement aléatoire est inférieure aux échelles de longueur caractéristiques du système. L'énergie associée à ce déplacement  $\delta U$  est donc inférieure à  $k_B T$  comme requis dans le MC-D.

### Principe d'ergodicité

Les scientifiques B. Alder et T. Wainwright lancèrent en 1957 les bases de la Dynamique Moléculaire, DM, actuelle [AW<sub>12</sub>]. Afin d'étudier la transition de phase solide-fluide, ils réalisèrent une simulation de sphères dures. Leurs résultats soulignèrent la convergence rapide vers la valeur d'équilibre des propriétés macroscopiques du système. À leur suite en 1957, les scientifiques W. Wood *et al.* réalisèrent une simulation Monte-Carlo [WP<sub>57</sub>, WJ<sub>57</sub>], basée sur le modèle précédent de Alder et Wainwright. Ils obtinrent les mêmes résultats.

L'équivalence des deux méthodes, DM et MC, souligna la véracité du **principe d'ergodicité**. Selon le principe d'ergodicité la moyenne de la variable aléatoire sur le temps est égale à la moyenne de la variable aléatoire sur l'ensemble des conformations. La méthode de DM permet d'extraire la moyenne temporelle, tandis que la méthode MC permet d'extraire la moyenne d'ensemble.

Le principe d'ergodicité pour un polymère est bien décrit par la dynamique du modèle de Rouse[DE<sub>88</sub>], voisin de notre approche. Ce modèle fut initialement établi en 1953 afin de rendre compte de la dynamique d'une chaîne libre, modélisée comme un collier de perle reliées par des ressorts. Il permet d'estimer le temps de relaxation de la chaîne de polymère entière qui varie en fonction du carré du degré de polymérisation,  $N$ .

### Lien avec le mouvement brownien

Remarquons que cette mesure de  $\langle dr_{rand}^2 \rangle$  peut être directement comparée à une mesure de mouvement brownien. Le mouvement brownien fut initialement introduit par le botaniste R. Brown afin de caractériser le mouvement continu et irrégulier auquel sont soumises des molécules de pollen dispersées dans de

l'eau [Bro28]. Ce mouvement stochastique est caractérisé par un processus de déplacement aléatoire en fonction du temps. Les premières explications théoriques du mouvement brownien furent données, indépendamment, par A. Einstein en 1905 [EE05] et M. Smoluchowski une année plus tard [vSo6].

Einstein entreprend une discussion à l'échelle microscopique de l'évolution temporelle du mouvement de particules dans un milieu. Ce mouvement erratique correspond à une alternance de mouvements libres des particules, et de chocs sans corrélations qui modifient leur trajectoire. Un tel processus est parfaitement illustré par une marche aléatoire, au vu de la direction aléatoire prise par la particule suite à un choc.

L'écart quadratique moyen définissant le déplacement d'une particule soumise au mouvement brownien, s'exprime ainsi :

$$\langle dx_{rand}^2 \rangle = \int x^2 \cdot P_x(\mathbf{x}, t) \cdot dx = 2D_{\text{iff}} \cdot \delta t \quad \text{en 1D} \quad (4.19)$$

Pour notre système subissant une diffusion équiprobable dans les 3 dimensions de l'espace, le déplacement quadratique moyen s'exprime donc par :

$$\Rightarrow \langle dr_{rand}^2 \rangle = 6D_{\text{iff}} \cdot \delta t \quad \text{en 3D} \quad (4.20)$$

### Durée d'un *pas* Monte-Carlo

L'égalisation des équations 4.17 et 4.20 nous permet d'évaluer le **pas de temps physique** associé à notre routine Monte-Carlo.

$$\Rightarrow \delta t = \frac{R_{dr_{rand}}^2}{10D_{\text{iff}}} \quad (4.21)$$

En fonction de la nature de l'objet qui a été choisi aléatoirement lors du pMC, le coefficient de diffusion diffère. Si l'objet correspond à la nanoparticule, sa diffusion sera plus lente que celle d'un monomère en vertu de la loi de Stokes.

### Acceptation ou refus des configurations

Si nous nous basons sur l'algorithme de Metropolis, il n'est pas nécessaire de connaître exactement la fonction de partition du système [MRR<sup>+</sup>53]. Les configurations sont acceptées ou refusées selon la loi de distribution de Maxwell-Boltzmann. Pour appliquer cela, la variation d'énergie interne du système,  $\Delta U$ , est calculée entre l'état initial et l'état final suite au déplacement aléatoire lors du pMC. Dans le cas où l'énergie du système a diminué lors de cette transition d'état,  $\Delta U < 0$ , la nouvelle configuration du système est acceptée et conservée. Dans le cas contraire, si l'énergie du système a augmenté,  $\Delta U > 0$ , la nouvelle configuration est acceptée avec une certaine probabilité, proportionnelle au facteur de Boltzmann.

$$P_r(i_{rand}, \mathbf{r}) = \exp\left(\frac{-\Delta U}{k_B T}\right) \quad (4.22)$$

Toute simulation Monte-Carlo Dynamique rend compte fidèlement de la dynamique du système tant que  $\Delta U \ll k_B T$ , ainsi le taux d'acceptation des nouvelles configurations aléatoires tend vers les 100%, [MTB<sup>+</sup>10].

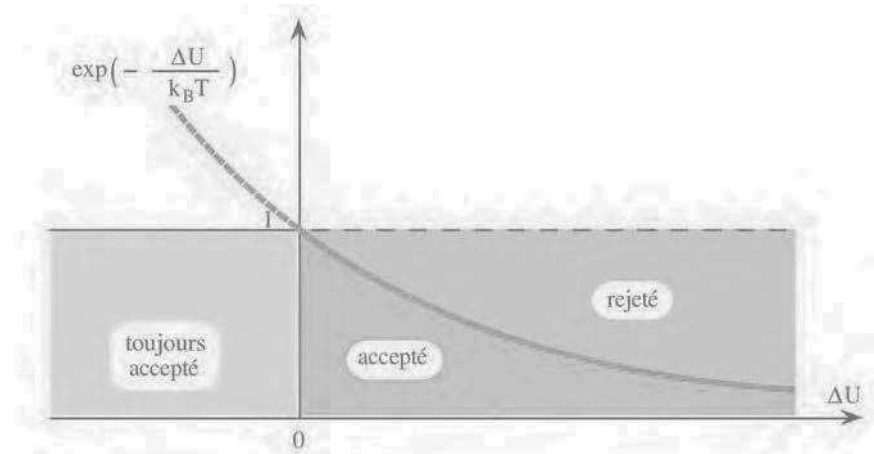


Fig. 4.6 – Représentation de la probabilité d'acceptation d'une configuration pondérée par le facteur de Boltzmann

L'acceptation du déplacement Monte-Carlo réalisé lors d'un pMC, suit donc les règles, illustrées en Fig. 4.6, suivantes :

- si  $\Delta U < 0$  le mouvement est toujours accepté
- si  $\Delta U > 0$  le mouvement est accepté selon une certaine probabilité, Eq. 4.22

#### Résumé d'un pas Monte-Carlo

Sur la durée physique  $\delta t$  est réalisé un pMC. Ce pas Monte-Carlo correspondant aux événements suivants :

- tirage d'un objet
- tirage de son déplacement
- vérification des contraintes de chaîne réelle
- calcul de la variation d'énergie interne au système si le déplacement avait lieu
- vérification de la probabilité d'accepter le déplacement
- déplacement, s'il est accepté.

Via l'algorithme de Monte-Carlo, nous avons un moyen simple et efficace d'obtenir des moyennes de grandeurs physiques dans un ensemble statistique. En effet, la partie conformationnelle de l'espace des états du système a été explorée. Dans cette méthode le temps n'est pas une variable explicite, laissant souvent l'image d'un processus *statique* qui ne rend pas compte de la dynamique du système. Les propriétés dynamiques sont cependant accessibles. Pour cela, il suffit d'intégrer et de prendre en compte convenablement la dimension temporelle.

#### 4.1.7 Évaluation du mouvement lors d'un balayage Monte-carlo, bMC

Afin de rendre compte des propriétés dynamiques du système il est maintenant nécessaire d'explicitier la notion de temporalité dans notre simulation. Pour cela, il faut définir et dissocier les différentes échelles de temps mises en jeu.

Le système est étudié à l'échelle de temps mésoscopique  $\delta t$ . Cette échelle de temps est petite par rapport au temps caractéristique d'observation du système, mais cependant, suffisamment grande pour pouvoir considérer que le mouvement brownien des particules sur deux intervalles de temps consécutifs sont indépendants. En d'autres termes, l'échelle de temps microscopique,  $\delta t_c$ , correspond à l'échelle de temps *collisionnel*, autrement nommée échelle de temps *moléculaire*. Ce temps *collisionnel* est donc petit devant  $\delta t$ . L'échelle de temps macroscopique correspond à l'échelle de temps caractérisant nos mesures d'observations.

Les échelles de temps mises en jeu dans nos simulation suivent donc la hiérarchie suivante :

$$T_{obs} \gg \delta t \gg \delta t_c \quad (4.23)$$

#### Durée d'une boucle Monte-Carlo

La durée d'un *balayage* Monte-Carlo correspond au temps au bout duquel tous les  $N$  objets constituant notre système ont eu la possibilité de bouger une fois en moyenne.

$$\Rightarrow t_{boucle} = N.\delta t \quad (4.24)$$

À chaque *balayage* Monte-Carlo, une séquence aléatoire d'états accessibles dans l'espace des configurations de notre système est générée. Ceci nous permet de faire le lien entre notre temps de simulation et la dynamique réelle. En effet, dans la dynamique réelle, tous les objets essaient de bouger simultanément. Donc un *pas* de temps physique correspond à  $N$  *pas* de temps MC.

#### Durée d'une simulation Monte-Carlo

La durée de la simulation, quantifiée en nombre de *balayage* Monte-Carlo effectuées, autrement dit la durée de notre mesure d'observation numérique, doit être suffisamment importante pour laisser le temps au système d'explorer l'ensemble des configurations. Là encore, nous devons laisser le temps au système de relaxer, afin d'obtenir ainsi un maximum d'informations sur les valeurs d'équilibre de notre complexe ADN/particule numérique.

De façon classique, la durée de nos simulations est de l'ordre de  $10^9$  à  $10^{10}$  étapes de *balayage* Monte-Carlo. Ceci nous permet de simuler jusqu'à l'équivalent de 2 s de variations de conformations du complexe ADN/particule, c'est à dire l'équivalent de 50 à 100 fois le temps de relaxation expérimental mesuré, qui est lui-même proche du temps de relaxation numérique [MTB<sup>+</sup>10]. Ainsi, le système a pu relaxer et explorer les différentes configurations accessibles. Cette durée de simulation a par ailleurs été confirmée par la cohérence entre les résultats de simulations réalisées avec différentes valeurs de nombres de *balayages* MC.

$$\Rightarrow T_{Simu} = 10^{10} \times N.\delta t \quad (4.25)$$

L'ensemble de ce processus nous permet d'explorer statistiquement l'espace des configurations de notre système ADN/particule.

### Fréquence d'observation

Nous voulons une fréquence d'observation du complexe numérique ADN/particule suffisamment grande pour que les positions du système soient, presque, décorrélées les unes des autres, mais suffisamment courte pour pouvoir suivre la dynamique du phénomène observé. Ainsi nous n'écrivons pas les données caractérisant le système sur l'ensemble des  $\delta t$ , ni même des  $t_{boucle}$ . Nous ajoutons un paramètre d'écriture des données  $t_{print}$  de l'ordre de quelques  $t_{boucle}$ .

L'évaluation d'une propriété d'équilibre du système est alors obtenue via une moyenne simple sur les configurations acceptées et écrites tous les  $t_{print}$ . Ainsi nous pouvons extraire la distance bout-à-bout de notre complexe simulé ADN/particule.

## 4.1.8 Exploitation des données

### Extraction des données

Afin de réaliser une comparaison directe entre les mesures expérimentales TPM et les simulations, nous extrayons la distance bout-à-bout de complexes numériques ADN/particule,  $\rho$ , Fig. 4.1.

La mesure de  $\rho$  correspond à la distance entre le centre du monomère initial  $i = 0$  ancré à la surface dont la position est fixée à  $(0, 0)$  et le centre de la nanoparticule indexée  $i = N - 1$ . Cette grandeur est donc à la projection en deux dimensions sur le plan (X,Y) du vecteur définissant la position du centre de la particule.

$$\langle R_{||,MC-D}^2 \rangle_{\boxed{2D}} = \sum_{i=t_{print}}^{N/t_{print}} \langle |\vec{r}_{N-1} - \vec{r}_0|^2 \rangle \quad (4.26)$$

$$\Rightarrow R_{||,MC-D} \boxed{2D} = \sqrt{\langle R_{||,MC-D}^2 \rangle_{\boxed{2D}}} \quad (4.27)$$

Dans un souci de simplicité nous noterons dans la suite  $R_{||,MC-D} \boxed{2D} = R_{||,MC-D}$ .

### Courbe de calibration

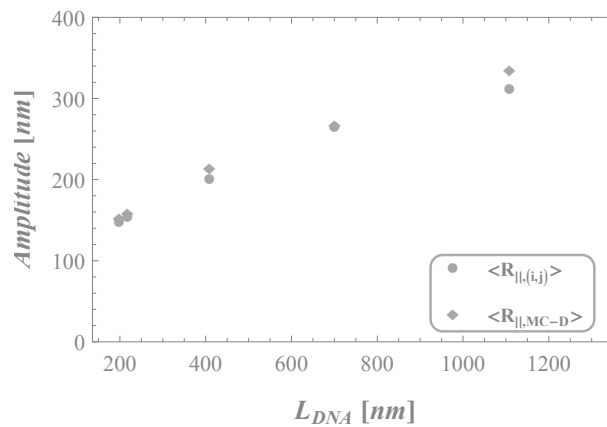


Fig. 4.7 – Courbe de calibration des amplitudes simulées,  $\langle R_{||,MC-D} \rangle$ , et expérimentales,  $\langle R_{||,(i,j)} \rangle$ . (Pour  $L_{ADN}=583,639,1201$  bp  $N=25+1$ , tandis que pour  $L_{ADN} = 2060,3259$  bp  $N=50+1$ ,  $R_{par} = 150$  nm pour toutes les simulations.

La validité de ce modèle de simulation via un algorithme de Monte-Carlo fut validée et démontrée en 2010 [MTB<sup>+</sup>10].

Les données brutes extraites de la simulation,  $R_{||,MC-D}$ , doivent être comparées aux données expérimentales  $R_{Exp||}$ , qui sont elles aussi en 2D. Autrement dit, l'amplitude expérimentale du mouvement en 2D,  $R_{TPM}$ , doit être corrigée de l'effet du moyennage vidéo, lié à l'écart entre le temps de relaxation des complexes ADN/particule,  $\tau_{||}$ , par rapport au temps d'acquisition du dispositif expérimental.

Il existe donc un bon accord entre les données simulées,  $R_{||,MC-D}$ , et les données expérimentales  $R_{Exp||}$ , Eq. 4.7.

### Influence de la discrétisation

Pour les ADN de 1201 et 3259 bp, il existe un léger décalage entre les deux estimations de l'amplitude des complexes ADN/particule représentées sur Fig. 4.7. Ceci est probablement induit par le niveau de discrétisation choisi dans les simulations, comme illustré sur Fig. 4.2.

En effet, le diamètre d'un des  $N = 25$  monomère décrivant l'ADN 1201 bp correspond à  $2a = 48$  bp dans la simulation. Nous sommes au-delà de la limite que nous nous sommes fixée, Eq. 4.5, afin de définir convenablement la taille des monomères. Ici, la taille d'un monomère n'est pas idéale. Il en est de même pour l'ADN 3259 bp dont les  $N = 50$  monomères ont leur taille définie égale à  $2a = 65$  bp dans la simulation.

Nous pouvons voir simplement l'influence de la discrétisation en calculant l'expression du modèle **WLC** en mode discret. Cette comparaison est beaucoup moins coûteuse en temps de calcul que si nous explorions cette question par simulation numérique de nos complexes ADN/particule sur une gamme de valeurs de  $N$ . L'effet, quand à lui, est du même ordre. Nous allons comparer nos données au modèle **WLC** ainsi qu'à ce dernier en régime discret, **WLC Discret** avec différentes valeurs de  $N$ .

Afin de comparer les modèles **WLC** et **WLC Discret**, à nos simulations, nous devons appliquer quelques corrections à leur expression classique. En effet, il faut :

- ajouter la contribution de la particule
- et enfin se restreindre au mouvement en 2 dimensions.

L'expression du modèle **WLC** avec la contribution de la particule selon la méthode minimale, le tout restreint en 2D, est la suivante :

$$\langle \mathbf{R}_{WLC}^2 \rangle = \frac{2}{3} \left[ 2L_p^2 \left( \frac{L}{L_p} - 1 + \exp\left(-\frac{L}{L_p}\right) \right) + R_{par}^2 \right] \quad (4.28)$$

De même, l'expression du modèle **WLC Discret** est la suivante :

$$\langle R_{WLC,D}^2 \rangle = \frac{2}{3} \left[ a^2 N W_N(v(\kappa_c)) + R_{par}^2 \right] \quad (4.29)$$

$$W_N(v(\kappa_c)) = \frac{1 + v(\kappa_c)}{1 - v(\kappa_c)} - \frac{2v(\kappa_c)}{N} \frac{1 - v(\kappa_c)^N}{(1 - v(\kappa_c))^2} \quad (4.30)$$

$$v(\kappa_c) = \coth(\kappa_c) - 1/\kappa_c \quad (4.31)$$

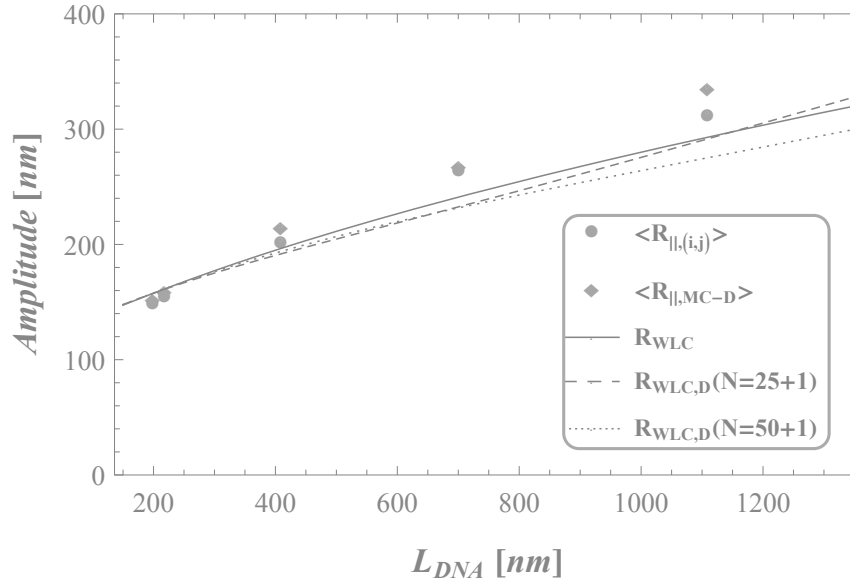


Fig. 4.8 – Courbe de calibration des amplitudes simulées,  $\langle R_{||,MC-D} \rangle$ , et expérimentales,  $\langle R_{||,(i,j)} \rangle$ . Pour  $L_{ADN} = 583, 639, 1201$  bp,  $N = 25 + 1$ , tandis que pour  $L_{ADN} = 2060, 3259$  bp,  $N = 50 + 1$ ;  $R_{par} = 150$  nm pour toutes les simulations. La courbe continue correspond au modèle **WLC**, la courbe discontinue correspond au modèle **WLC Discret** pour  $N = 25 + 1$  tandis que la courbe en pointillés correspond à  $N = 50 + 1$

Aux grands  $L_{ADN}$ , l'effet de la discrétisation sur  $R_{WLC,D}$  est nettement visible, Eq. 4.8. Une discrétisation en  $N$  plus convenable, c'est-à-dire  $N$  augmente, diminue la valeur de l'amplitude du mouvement. Ceci montre bien que l'écart pour les ADN de longueur 1201 et 3259 bp entre nos simulations et l'expérience est, entre autre, dû au défaut dans la valeur de  $N$  utilisée.

De plus, nous remarquons que plus la longueur de l'ADN augmente plus l'écart est important entre nos données, numériques et expérimentales, par rapport aux données des modèles théoriques. En effet, comme nous l'avons vu précédemment, les modèles **WLC** ainsi que **WLC Discret** ne prennent aucunement l'effet de volume exclu en compte. Or, cet effet de volume exclu est plus important pour les ADN de grande taille. Il en résulte que l'erreur induite par l'utilisation des modèles théorique afin d'estimer les propriétés des chaînes réelles est plus importante sur les ADN long. Les modèles **WLC** ainsi que **WLC Discret** ne prennent non plus pas en compte la présence de la surface solide. Cela justifie pleinement l'utilisation d'un modèle numérique pour rendre compte des effets fins que nous étudierons par la suite.



### Influence du volume exclu

Intéressons-nous d'un peu plus près à la contribution du volume exclu dans l'amplitude du mouvement. Pour cela, nous faisons varier la taille du volume exclu imposé dans nos simulations numériques sur deux longueurs d'ADN différentes. Afin de couvrir la gamme de comportement du polymère nous nous focalisons sur l'ADN de 575 bp, dont l'effet de volume exclu à théoriquement peu d'influence, et sur l'ADN de 2060 bp, plus sensible aux effets de gonflement de la chaîne. Ceci fait écho à notre première prospection des effets de volume exclu en fonction du modèle théorique utilisé, Chapitre 2.

Le volume exclu maximal est égal au volume occupé par un monomère. Nous diminuons progressivement le volume d'exclusion proportionnellement à  $a_{\text{mon}}$  pour les deux ADN. Nous caractérisons la valeur du volume exclu via le rayon de la sphère d'exclusion ainsi que de du rayon de la zone d'exclusions du potentiel Lennard-Jones, renseigné dans la Table 4.1.

$V_{\text{exc}}$ nm	$L_{\text{ADN}} = 583$ bp   nm		$L_{\text{ADN}} = 2060$ bp   nm	
$a_{\text{mon}}$	12	4	21	7
$\frac{6}{5}a_{\text{mon}}$	9	3	15	5
$\frac{4}{5}a_{\text{mon}}$	6	2	10	4
$\frac{3}{5}a_{\text{mon}}$	4	1	8	3
$\frac{2}{5}a_{\text{mon}}$	3	1	5	2
$\frac{1}{5}a_{\text{mon}}$	1	0.5	3	1
0	0	0	0	0

TABLE 4.1 – Valeur du volume exclu

De plus, nous testons l'effet de la taille du volume exclu à différentes longueurs de persistance :  $L_p^{(1)} = 147 \text{ bp} = 50 \text{ nm}$  et  $L_p^{(2)} = 183 \text{ bp} = 62 \text{ nm}$ .

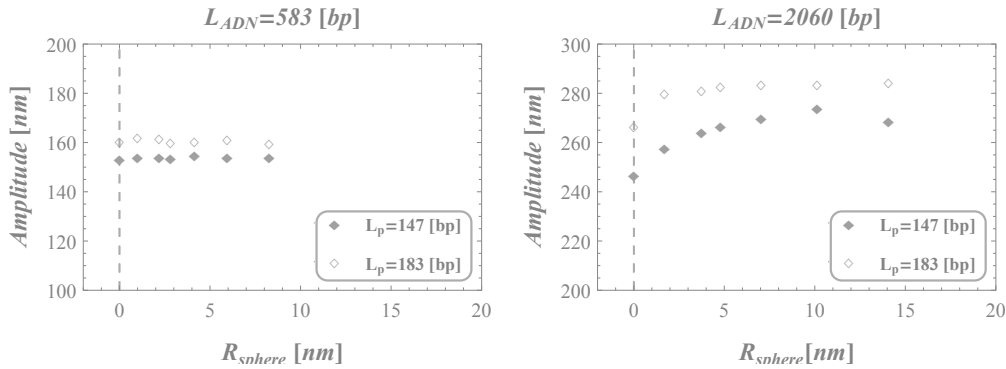


Fig. 4.9 – Influences du volume exclu sur l'estimation simulée,  $\langle R_{||,MC-D} \rangle$  Pour  $L_{\text{ADN}} = 583 \text{ bp}$ ,  $N = 25 + 1$  et pour  $L_{\text{ADN}} = 2060 \text{ bp}$ ,  $N = 50 + 1$ ,  $R_{\text{par}} = 150 \text{ nm}$  pour toutes les simulations.  $L_p^{(1)}$ , données en symboles pleins, et  $L_p^{(2)}$ , données en symboles vides

Ceci est une nette illustration de l'effet de gonflement de la chaîne dû au volume exclu, Fig. 4.10. Les polymères de petite taille restent insensibles à cet effet. Par contre, l'effet induit par la présence de volume exclu est clairement visible sur  $L_{\text{ADN}} = 2060 \text{ bp}$ . Cet effet devient d'autant plus important que l'ADN est long, la probabilité de rencontre de la chaîne avec elle-même augmente avec la longueur de la molécule.

De plus, l'amplitude de cet effet est plus importante pour  $L_p^{(1)} = 147$  bp que pour  $L_p^{(2)}$ . En effet, plus la longueur de persistance augmente, plus l'ADN devient rigide, donc plus la probabilité de rencontre de la chaîne avec elle-même diminue. Lors de changements de configuration de la chaîne, l'effet de volume exclu contribue donc moins souvent lorsque la longueur de persistance augmente.

Afin d'aller plus loin et comparer nos données numériques aux modèles théoriques, nous réalisons un ensemble de simulation sur une gamme de longueur de persistance. Nous compilons ces données avec les données Fig.4.10 suivantes.

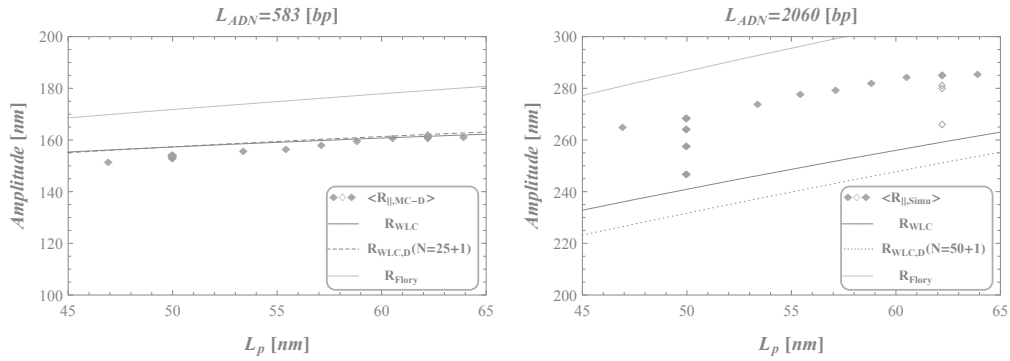


Fig. 4.10 – Influence de la longueur de persistance et du volume exclu sur l'estimation simulée,  $\langle R_{||,MC-D} \rangle$  Pour  $L_{ADN} = 583$  bp,  $N = 25 + 1$  et pour  $L_{ADN} = 2060$  bp,  $N = 50 + 1$ ,  $R_{par} = 150$  nm pour toutes les simulations.  $L_p^{(1)}$ , données en symboles pleins, et  $L_p^{(2)}$ , données en symboles vides. La courbe continue correspond au modèle **WLC**, la verte au modèle de **Flory**.

Comme attendu, l'amplitude du mouvement dépend de la longueur de persistance. Pour l'ADN court, l'effet du volume exclu reste indistinguable. De plus, l'accord entre les simulations et le modèle reste proche. L'écart entre ces deux estimations est plus important,  $\sim 5$  nm, lorsque  $L_p$  diminue. Dans cette situation, la valeur de la longueur de persistance se rapproche de la taille des monomères décrivant l'ADN, ce qui peut contribuer à cet écart.

L'évaluation de l'amplitude du mouvement des complexes ADN/particule simulés pour l'ADN 2060 bp, n'est pas superposable avec l'approche théorique **WLC**. Cette différence entre les deux modèles diminue, passant de  $\sim 20$  nm à  $\sim 5$  nm, lorsque le volume exclu inclus dans la simulation diminue. Lorsque ce dernier est mis à zéro, afin d'être comparable au modèle **WLC** qui ne prend pas en compte cet effet, l'écart entre simulation et théorie diminue.

La prise en compte de l'effet de volume exclu est un paramètre important afin de rendre compte convenablement de notre géométrie expérimentale. Les modèles théoriques classiques ne permettent pas de pleinement rendre compte de l'influence de ce critère, que ce soit le modèle **WLC** ou celui de **Flory**.

#### 4.1.9 Temps de corrélation

Le temps de corrélation du système numérique peut être calculé, grâce aux mêmes équations, définissant les fonctions de corrélation, détaillées précédemment dans les parties 3.7.5 et 3.7.5. Pour les simulations, aucune correction liée à l'effet de moyennage vidéo n'est, bien sûr, nécessaire.

Afin de tester la justesse de notre simulation MC Dynamique, les temps caractéristiques des systèmes simulés sont comparés aux temps de relaxation expérimentaux moyennés sur typiquement une cinquantaines de trajectoires [MTB<sup>+</sup>10].

$L_{DNA}$ , bp	Temps de corrélation du système : $\tau_{  }$ , ms	
	Données TPM	Données MC-D
2080	57.4	48.7
1500	42.7	37.0
798	29.7	20.6
401	20.0	12.2

Nous faisons l’hypothèse que la dynamique des simulations MD-D rend correctement compte de la dynamique réelle que nous avons pu mesurer expérimentalement.

De plus, sur l’ensemble des systèmes que nous avons explorés via ces simulation le taux d’acceptation des conformations,  $\tau_{acc}$  varie de de 75% à 85%.

Tout ceci valide la fiabilité de nos simulations MC-D par l’exploration correcte de l’espace des énergies de configurations associées au système physique ainsi que de la dynamique du système qui rend fidèlement compte de la dynamique réelle.

En moyenne une simulation de ce type, **Monte-Carlo Dynamique**, dure de 4 jours à 1 semaine sur un seul processeur. Or, lors de notre étude, les informations que nous cherchons ne sont pas toutes de l’ordre de la dynamique du système. Afin de rendre compte rapidement des propriétés statistiques intrinsèques aux complexes ADN/particule que nous étudions expérimentalement, nous avons élaboré une simulation dite d’**échantillonnage exact**. Cette nouvelle simulation est basée sur le même principe que cette précédente méthode Monte-Carlo, mais rend compte des propriétés conformationnelles purement *statiques* du système.

## 4.2 SIMULATION PAR ÉCHANTILLONNAGE EXACT (EE)

Les conformations accessibles au complexe ADN/particule peuvent aussi être calculées numériquement par une simulation d’**échantillonnage exact** basée sur [SNPo6].

Le modèle d’**échantillonnage exact** via tirage aléatoire, EE, suit la même définition du système que dans le modèle MC-D précédant. Les modélisations du monomère, 4.1.1, ainsi que de la particule, 4.1.2, sont construites similairement.

La conformation du système n’est maintenant, plus explicitement régie par son énergie interne. Précédemment, cette dernière nous permettait de contraindre sa configuration au domaine énergétiquement accessible. De plus, nous ne cherchons pas à rendre compte de la dynamique du système. Nous pouvons donc définir la conformation de ce dernier via la fonction de distribution d’équilibre des angles  $\theta$  et  $\phi$  accessibles entre chaque maillon du système, comme précédemment défini Fig. 2.8.

Seuls les angles  $\theta$  et  $\phi$  entre chaque objet de notre système sont tirés aléatoirement. Ainsi, seule la distribution de probabilité des angles  $\theta$  et  $\phi$  rend compte de la conformation.

#### 4.2.1 Simulation des contraintes expérimentales

##### Sphères dures et volume exclu

Chaque monomère est toujours considéré comme une sphère dure de rayon  $a$ , de même pour la nanoparticule de rayon  $R_{par}$ .

Nous appliquons les conditions de volume d'exclusion sur les objets définis dans notre système. Pour cela, nous calculons la distance existant entre le centre de l'objet  $i$  soit  $(x_i, y_i, z_i)$ , et le centre de l'objet  $j$  soit  $(x_j, y_j, z_j)$ . Si cette distance est inférieure à la distance limite d'exclusion, alors la conformation est refusée. Dans cette simulation nous n'utilisons pas le potentiel de Lennard-Jones.

Les valeurs des distances entre les centres de deux objets de notre système doivent être supérieures à  $R_{min}$  :

- pour le cas monomère-monomère :  $R_{min} = 2a$  ;
- pour le cas monomère-particule :  $R_{min} = a + R_{par}$ .

##### Murs durs

Afin de rendre compte de la présence d'une surface solide dans notre dispositif expérimental, nous appliquons les mêmes conditions de murs durs que précédemment.

La distance entre la surface et le centre de l'objet doit rester supérieure à  $z_{max}$  :

- pour les monomères :  $z_{max} = a$  ;
- pour la particule :  $z_{max} = R_{par}$ .

#### 4.2.2 Évaluation des conformations du système

La distribution de probabilité d'équilibre des angles  $\theta$  et  $\phi$ , définissant la position de chaque monomère par rapport au précédent, dépend de la nature de l'objet.

##### Monomère et segment initial

La position du premier monomère est, comme précédemment, fixée à zéro dans le plan  $(X, Y)$  et placé en  $a$  selon l'axe des  $Z$ . Afin de rendre compte de la présence du mur, la distribution du vecteur défini par les positions des centres du monomère initial,  $i = 0$ , et du suivant,  $i = 1$  doit être contrainte. Ce qui implique que la distribution de l'angle  $\theta$  doit être tronqué. Ainsi, le tirage aléatoire des paramètres angulaires pour le monomère  $i = 1$  est uniformément inclus dans les domaines suivants :

- $\phi \in [0 : 2\pi]$
- $\theta \in [0 : \pi/2]$

### Autres monomères

La position des monomères suivants doit satisfaire les contraintes induites par le module de courbe du polymère. Dans la simulation, le module de courbure est assimilé à :

$$\kappa_c = \frac{L_p k_B T}{2a} \quad (4.32)$$

Afin de rendre compte de la flexibilité du polymère, est distribué à l'équilibre selon la loi de Boltzmann. Nous montrons que, suite au tirage d'un nombre aléatoire  $i_{rand} \in [0 : 1]$ , l'expression de  $\cos \theta$  est alors définie par :

$$\cos \theta = \frac{\log [\exp(-\kappa_c) + i_{rand}(\exp(\kappa_c) - \exp(-\kappa_c))]}{\kappa_c} \quad (4.33)$$

Le tirage aléatoire du paramètre angulaire  $\phi$  est uniformément distribué dans le domaine :  $\phi \in [0 : 2\pi]$ .

### Nanoparticule

Nous considérons la nanoparticule comme un dernier joint libre. Ainsi les tirages aléatoires des paramètres angulaires sont distribués dans les domaines suivants :

$$\begin{aligned} - \phi &\in [0 : 2\pi] \\ - \theta &\in [0 : \pi] \end{aligned}$$

uniformement sur l'angle solide  $4\pi$ .

#### 4.2.3 Exploitation des données

Le point fort de cette simulation EE est le gain en temps de calcul. La durée nécessaire pour une simulation est maintenant inférieure à la demi-journée. En effet, il n'est plus nécessaire de laisser le temps au système de relaxer, ni de le laisser relaxer entre les deux mesures.

Le coût limité en temps de calcul des simulations EE, nous permet d'augmenter la discrétisation du modèle. Ainsi, nous réalisons nos simulations avec la valeur par défaut de la taille des monomères mise égale à  $a_{mon} = 6$  bp. Ceci est de l'ordre du diamètre de l'ADN dans les conditions physiologiques de température et de conditions salines.

En revanche cette technique ne donne plus accès à la dynamique du système, tel que le temps de relaxation du système.

### Extraction des données

La procédure d'extraction de la distance bout-à-bout des complexes numériques ADN/particule est en tout point identique à celle utilisée suite aux simulations MC-D précédentes.

### Courbe de calibration

L'accord entre les données de la simulation EE et les données expérimentales est très bon, de même avec celles de la simulation MC-D, Fig. 4.11.

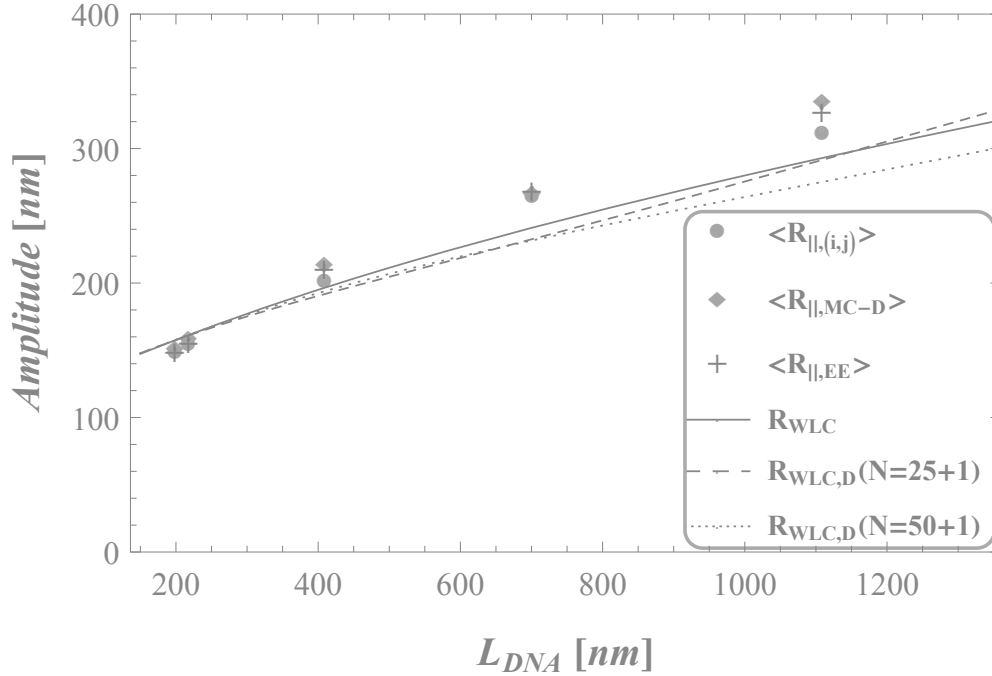


Fig. 4.11 – Amplitudes simulées,  $\langle R_{||,EE} \rangle$ , symbole +, avec  $L_{ADN} = 583, 639, 1201, 2060, 3259$  bp avec respectivement  $N = (49 + 1), (53 + 1), (100 + 1), (172 + 1), (272 + 1)$  afin de respecter  $a_{mon} = 6$ bp. Amplitudes simulées via MC-D,  $\langle R_{||,MC-D} \rangle$ , symbole  $\diamond$ , pour  $L_{ADN} = 583, 639, 1201$  bp  $N = 25 + 1$ , tandis que pour  $L_{ADN} = 2060, 3259$  bp  $N = 50 + 1$ .  $R_{par} = 150$  nm pour toutes les simulations. Données expérimentales,  $\langle R_{||,i,j} \rangle$ . La courbe continue correspond au modèle WLC, la courbe discontinue correspond au modèle WLC Discret pour  $N = 25 + 1$  tandis que la courbe en pointillés correspond à  $N = 50 + 1$

La figure 4.11, souligne le bon accord entre les simulations EE et MC-D, qui, toutes deux sont en accord avec les données expérimentales. Pour l'ADN 3259 bp, la valeur extraite par EE diminue par rapport à celle via MC-D, se rapprochant des données expérimentales. Ceci est cohérent avec l'influence du niveau de discrétisation du polymère dans la simulation. En effet, comme discuté plus haut pour le modèle **WLC Discret**, les valeurs obtenues avec  $N = 50 + 1$  sont plus basses que lorsque  $N = 25 + 1$ .

En nous focalisant sur l'ADN 2060, nous prospectons les variations d'amplitudes induites par la variation de rigidité du polymère, Fig. 4.12. Une série de simulations est obtenue avec les mêmes paramètres que celles ci-dessus, sur une gamme de longueur de persistance variant comme précédemment avec MC-D. Ces simulations sont réalisées avec un volume exclu défini par un rayon égal à  $a_{mon} = 6$  bp soit 2 nm environ.

Un écart apparaît pour les longueurs de persistance grandes entre les données EE et MC-D. Notons que ce type de modifications sur la longueur de persistance influe sur le temps de relaxation du système. En effet, une augmentation de la rigidité du polymère ralentit sa dynamique. Il est donc nécessaire de prendre en compte cet effet et donc d'ajuster le temps de calcul dans les simulations MC-D. Les simulations en échantillonnage exacts EE, ne sont pas affectées par ce type de contraintes. Si nous nous intéressons seulement aux propriétés statistiques conformationnelles du polymère, cet outil est très adapté.

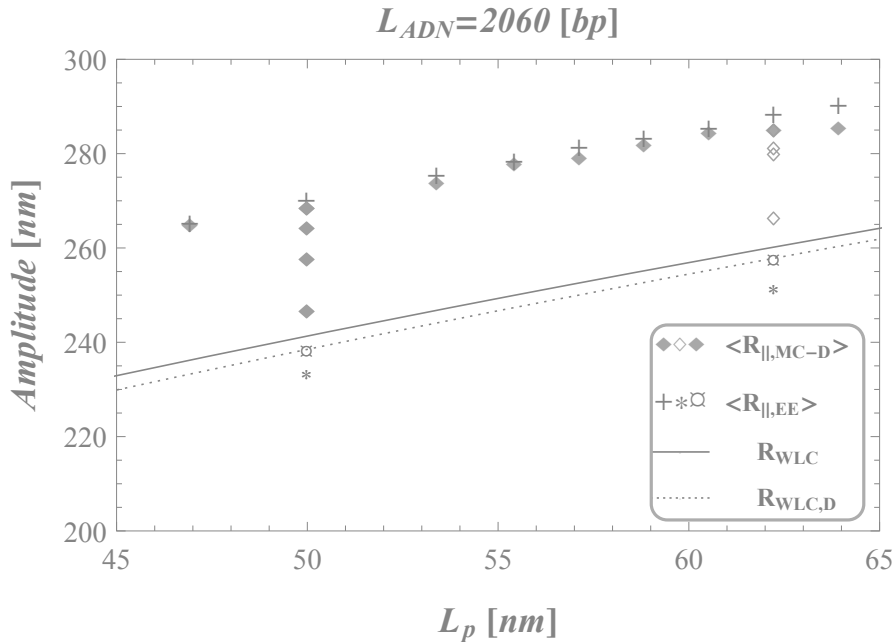


Fig. 4.12 – Courbe de calibration des amplitudes simulées,  $\langle R_{||,EE} \rangle$ , symbole +, avec  $L_{ADN} = 2060$  bp avec  $N = (172 + 1)$  afin de respecter  $a_{mon} = 6$  bp. Amplitudes simulées via MC-D,  $\langle R_{||,MC-D} \rangle$ , symbole  $\diamond$ , pour  $L_{ADN} = 2060$  bp avec  $N = 50 + 1$ .  $R_{par} = 150$  nm pour toutes les simulations. La courbe continue correspond au modèle WLC, la courbe discontinue correspond au modèle WLC Discret pour  $N = 25 + 1$  tandis que la courbe en pointillée correspond à  $N = 50 + 1$

Maintenant, nous prospectons l'effet de volume exclu existant sur les différentes parties du système. Pour cela nous nous sommes focalisés sur des simulations pour  $L_p = 147$  et  $183$  bp, où nous avons testé les deux types d'effets suivants :

- de volume exclu
- de mur dur

Sans volume d'exclusion, l'interpénétration de la chaîne avec elle-même est possible. De plus, sans mur, même si le monomère initial est fixé en  $(0, 0)$  dans le plan  $(X, Y)$ , le complexe ADN/particule peut explorer l'ensemble de l'espace correspondant à une sphère entière et non plus seulement une demi-sphère, comme dans l'expérience.

La figure 4.12 présente quelques simulations où nous avons testé l'influence sur le mouvement des effets ci-dessus en appliquant les conditions suivantes :

- pas de volume exclu et pas de mur (Simulation indexée [000], symboles  $\diamond$ )
- pas de volume exclu (Simulation indexée [011], symboles \*)

L'effet dû au volume exclu est pleinement pris en compte dans le modèle EE. Lorsque que nous supprimons les effets de volume d'exclusions induits par la présence du mur et par la non interpénétration de la chaîne, les valeurs de l'amplitude diminuent fortement.

Intéressons-nous à l'effet induit par la présence de la surface solide. Pour cela nous supprimons la contrainte due au mur dur. Nous pouvons constater un accord parfait entre les simulations EE sans volume exclu et sans mur avec le modèle **WLC-Discret**. En effet, ce modèle théorique ne prend pas en compte ces deux effets dans sa description analytique. Ceci nous permet de vérifier que ce modèle de simulation EE, rend compte exactement des conformations accessibles à la chaîne de polymère dans cette configuration où les calculs analytiques exacts sont possibles.

### Acceptation ou refus des configurations

Ce modèle de simulation par échantillonnage exact possède un plus grand taux de refus des configurations testées. En effet, contrairement aux modèles MC-D dont l'espace des configurations suit la loi de distribution de Boltzmann, ici toutes les configurations possibles sont testées. Ainsi, un grand nombre de configurations énergétiquement acceptables sont testées avant d'être refusées car elles conduisent à au moins une interpénétration des objets du système.

Intéressons-nous au taux d'acceptation des configurations associées à ces simulation EE en fonction des paramètres de volume exclus que nous appliquons.

Volume exclus			$L_{ADN} = 583$		$L_{ADN} = 2060$		Symbole Fig. 4.12
Mur	Chaîne	Particule	$\tau_{acc}$ %	$R_{  ,EE}$ nm	$\tau_{acc}$ %	$R_{  ,EE}$ nm	
0	0	0	100	158	100	238	$\square$
0	0	1	25	191	14	285	
0	1	1	24	191	13	287	*
0	1	0	99	158	94	241	
1	1	0	12	137	14	235	
1	0	0	12	137	15	234	
1	0	1	5	151	3	269	
1	1	1	5	151	2	271	+

TABLE 4.2 – Données simulations EE, variation du taux d'acceptance des conformations du système, et de l'amplitude, en fonction des volumes exclus pris en compte dans la simulation. Avec  $L_{ADN} = 583, 2060$  bp avec respectivement  $N = (49 + 1), (172 + 1)$  afin de respecter  $a_{mon} = 6bp$ ,  $L_p = 147$  bp et  $R_{par} = 150$  nm pour toutes les simulations.

Le taux d'acceptation des conformation du système diminue d'autant plus que l'espace des conformations accessible à la chaîne est contraint, c'est-à-dire en fonction de la prise en compte et de l'importance des volumes exclus mis en jeu.



## CONCLUSION DU CHAPITRE

Nous avons pu constater l'écart entre nos mesures TPM et les modèles classiques de la physique des polymères. Afin de comprendre au mieux ces différences et surtout les paramètres influant sur notre système ADN/particule, nous avons construit un modèle numérique basé sur un algorithme Monte-Carlo. La comparaison entre tous ces outils nous a permis de souligner les paramètres cruciaux influant sur les propriétés statistiques du système.

La précision nécessaire à l'étude des propriétés statistiques du polymère dépend de la finesse des processus que nous cherchons à étudier. Ainsi, lorsque nous cherchons à remonter aux propriétés physiques de la molécule d'ADN avec une précision au nanomètre près, les outils que nous utilisons doivent être des plus rigoureux et pertinents.

Afin d'obtenir une estimation de l'erreur induite par l'extraction des propriétés physiques de l'ADN, nous réalisons une estimation inversée de la longueur de persistance associée à une simulation Monte-Carlo.

Soit un ADN, de longueur de persistance  $L_{p,Simu} = 50$  nm, d'une longueur de 2060 bp simulée par  $N = 50 + 1$  monomères auquel est accroché à son extrémité libre une nanoparticule de rayon  $R_{par} = 150$  nm. La mesure de son amplitude du mouvement,  $R_{||,MC-D}$ , projection en 2 dimensions de la distance bout-à-bout du système est égale à 269 nm. En utilisant la méthode minimale, 3.8.2, pour extraire la distance bout-à-bout de l'ADN, puis le modèle **WLC**, 2.21, afin d'extraire la longueur de persistance, nous obtenons  $L_p = 68$  nm.

Sur le même principe, nous simulons l'ADN 583 bp via  $N = 25 + 1$  monomères. L'amplitude du mouvement mesurée, suite à la simulation, est égale à 154 nm. Nous obtenons, lors de l'extraction de la longueur de persistance  $L_p = 41$  nm, alors que la valeur numérique imposée dans la simulation est égale à  $L_{p,Simu} = 50$  nm. La valeur de la longueur de persistance *inverse*, diffère nettement de la valeur de longueur de persistance incluse dans la simulation.

Le désaccord entre ces deux valeurs de longueurs de persistance correspond à une erreur relative allant de 20 à 30%, en fonction de la longueur de l'ADN étudié. Cet écart est lié à différentes erreurs. Ces erreurs proviennent, essentiellement, de la combinaison de l'erreur existante sur l'estimation de la valeur de  $R_{||,MC-D}$  due à notre modèle de simulation ainsi que de l'erreur induite par la méthode d'extraction de  $L_p$ . Or, au vu de l'accord entre les valeurs des amplitudes du mouvement obtenues via les simulations et via l'expérience, l'erreur induite par notre modèle de simulation n'est pas si importante, même si elle reste non nulle. De ce fait, le désaccord entre  $L_{p,Simu}$  et  $L_p$  extraite, est essentiellement du à la méthode d'extraction. Tout ceci souligne la complexité de l'extraction des propriétés physiques de la molécule d'ADN. Nous en reparlerons au chapitre suivant.



# OBJECTIFS DE LA THÈSE

Comprendre les effets de l'environnement sur la mécanique de l'ADN est essentiel pour mieux appréhender ce qui se produit à l'intérieur du noyau cellulaire, notamment lors de processus biologiques tels que la transcription ou la réplication. En effet, il a été constaté expérimentalement que l'ADN est un biopolymère qui peut se rigidifier, se courber abruptement, voire se dénaturer en fonction de sa séquence, mais aussi en fonction des conditions physico-chimiques de son environnement. De tels événements modifient radicalement **la flexibilité et l'élasticité de l'ADN**, autrement dit la **longueur de persistance** de ce biopolymère. La détection de l'impact sur les propriétés mécaniques et physiques de la molécule d'ADN par de tels **effets globaux ou d'événements locaux**, est cruciale pour la compréhension de la physique de cette molécule.

En effet, encore aujourd'hui de nombreuses questions se posent quant à la flexibilité de l'ADN. De nombreux processus biologiques *in vivo* ne peuvent être simplement expliqués par les modèles classiques de physique des polymères tels que **Freely-Jointed-Chain** ou **Worm-Like chain**. De plus, nous avons pu souligner les limites de ces modèles afin d'extraire les propriétés physiques intrinsèques à la molécule d'ADN.

Dans ce cadre-là, nous utilisons et développons des outils à l'échelle de la molécule unique afin de quantifier et de prospecter, à la fois, expérimentalement et théoriquement, l'impact de ces modifications globales ou locales. Pour cela nous utilisons les capacités des méthodes de suivi de molécule unique via la technique expérimentale de **Tethered Particle Motion**. L'avantage majeur de cette technique réside dans l'absence de forces extérieures appliquées au système, ne contraignant ainsi pas significativement les degrés de liberté conformationnelle du polymère. Touchant aux limites des modèles analytiques, nous utiliserons un **modèle numérique de physique statistique mésoscopique**, basé sur un algorithme de Monte Carlo, afin de résoudre le problème inverse. Nous combinons donc approche expérimentale et approche numérique basée sur **les modèles de physique des polymères**. En effet, l'outil numérique est central dans notre approche, car la connaissance du modèle sous-jacent est déterminante dans l'interprétation des effets fins que nous cherchons à mettre en exergue.

Dans cette première partie consacrée aux résultats, nous allons nous intéresser à l'impact de l'environnement physico-chimique défini par la force ionique du milieu, sur la mécanique d'une molécule ADN. Quelle est l'influence de la force ionique sur la longueur de persistance du biopolymère qu'est la molécule d'ADN? Cette question fut soulevée depuis de nombreuses décennies. Dans les années 80, une longue série d'études fut réalisée et elles restent aujourd'hui

toujours sujettes à controverse. En effet, les prévisions théoriques du modèle Odijk-Skolnick-Fixman, suivies de près par le modèle de condensation de contre-ions de Manning, s'écartent des données expérimentales de la littérature. Nous cherchons donc à acquérir des données expérimentales sur une gamme de forces ioniques étendue et sondée extensivement. Tout d'abord, pouvons-nous quantifier une variation des propriétés physiques de la molécule d'ADN en fonction de la force ionique induite par la présence de contre-ions en solution ? Quel est le profil de dépendance de cette dernière sur les propriétés physiques et plus particulièrement sur la longueur de persistance de l'ADN ? L'influence de la force ionique sur les propriétés physiques de l'ADN est-elle dépendante de la nature des contre-ions présents en solution, monovalents avec  $Na^+$  et divalents avec  $Mg^{2+}$  ? Pouvons-nous comprendre la cause des écarts entre théorie et expérience ? Ainsi, nous est-il possible de contribuer à une meilleure compréhension de la prise en compte de l'effet induit par une variation de la force ionique ? Nous nous efforcerons de répondre à l'ensemble de ces questions dans ce premier chapitre de résultats.

Dans le deuxième chapitre de cette partie consacrée aux résultats, nous nous intéresserons à l'influence d'effets locaux induits par la présence de courbure intrinsèque due à la séquence même de l'ADN, ou encore par la formation d'une bulle de dénaturation. Dès les années 80, des séquences spécifiques existant *in vivo* furent identifiées par leur particularité d'induire une courbure intrinsèque à la séquence même de la molécule d'ADN. Ceci est induit par la succession des acides nucléiques les constituant, autrement dit par la séquence même de l'ADN. Ces séquences courtes, de l'ordre d'une dizaine de paires de bases, auxquelles nous nous intéressons se nomment A-tract. Nous est-il possible de détecter une courbure dans la molécule d'ADN utilisée en TPM ? Les contraintes mécaniques induites par ces séquences courtes ont-elles un effet tout le long de l'ADN et sont-elles détectables sur une taille de molécule importante ? Nous est-il possible de détecter et quantifier une variation d'amplitude ainsi que la courbure induite par la présence d'un A-tract ?

Après avoir examiné ces questions, nous nous intéresserons à l'influence de la température sur la molécule d'ADN. Nous est-il possible de distinguer le régime où le double brins d'ADN est encore hybridé du régime où il existe des parties déshybridées dans la molécule d'ADN, c'est à dire où il existe une bulle de dénaturation ? Quelle est la dépendance des propriétés physiques, et plus particulièrement de la longueur de persistance de l'ADN, en fonction de la température ? Existe-il une concordance entre les résultats expérimentaux et les modèles théoriques ? Nous explorerons l'ensemble de ces questions dans le troisième chapitre de résultats.

## BIBLIOGRAPHIE

- [AW12] B Alder and T Wainwright. Phase Transition for a Hard Sphere System. *Journal of Chemical Physics*, 27(5) :1208–1211, 2012.
- [Bro28] Robert Brown. XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies, volume 4. 1828.
- [DDPM13] Anil Kumar Dasanna, Nicolas Destainville, John Palmeri, and Manoel Manghi. Slow closure of denaturation bubbles in DNA : Twist matters. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 87(5) :052703, may 2013.
- [DE88] M. Doi and S. F. Edwards. *The theory of polymer dynamics*, volume 73. oxford university press, 1988.
- [Dor91] Robert Landick Dorothy A. Schafer, Jeff Gelles, Micheal P. Sheetz. Transcription by single molecules of RNA polymerase observed by light microscopy. *Nature*, 352 :444–447, 1991.
- [DS06] Nicolas Destainville and Laurence Salomé. Quantification and correction of systematic errors due to detector time-averaging in single-molecule tracking experiments. *Biophysical journal*, 90(2) :L17–L19, jan 2006.
- [DSC<sup>+</sup>14] Cheikh Tidiane Diagne, Maya Salhi, Estelle Crozat, Laurence Salomé, Francois Cornet, Philippe Rousseau, and Catherine Tardin. TPM analyses reveal that FtsK contributes both to the assembly and the activation of the XerCD-dif recombination synapse. *Nucleic Acids Research*, 42(3) :1721–1732, feb 2014.
- [DZMF11] David Dunlap, Chiara Zurla, Carlo Manzo, and Laura Finzi. Probing DNA topology using tethered particle motion. *Methods in Molecular Biology*, 783 :295–313, jan 2011.
- [EE05] a. Einstein and a. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 322(8) :549–560, 1905.
- [GH68] H B Gray and J E Hearst. Flexibility of native DNA from the sedimentation behavior as a function of molecular weight and temperature. *Journal of molecular biology*, 35(1) :111–129, jul 1968.
- [GKV11] Stephanie Geggier, Alexander Kotlyar, and Alexander Vologodskii. Temperature dependence of DNA persistence length. *Nucleic Acids Research*, 39(4) :1419–1426, mar 2011.
- [Hag88] P J Hagerman. Flexibility of DNA. *Annual review of biophysics and biophysical chemistry*, 17 :265–286, 1988.
- [Hir76] T Hirschfeld. Optical microscopic observation of single small molecules. *Applied optics*, 15(12) :2965–2966, dec 1976.

- [JLESESo6] Prashant K. Jain, Kyeong Seok Lee, Ivan H. El-Sayed, and Mostafa a. El-Sayed. Calculated absorption and scattering properties of gold nanoparticles of different size, shape, and composition : Applications in biological imaging and biomedicine. *Journal of Physical Chemistry B*, 110(14) :7238–7248, apr 2006.
- [Jon24] J. E. Jones. On the Determination of Molecular Fields. I. From the Variation of the Viscosity of a Gas with Temperature. *Proceedings of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 106(738) :441–462, oct 1924.
- [LAR<sup>+</sup>12] Timothée Lionnet, Jean François Allemand, Andrey Revyakin, Terence R. Strick, Omar a. Saleh, David Bensimon, and Vincent Croquette. Single-molecule studies using magnetic traps. *Cold Spring Harbor Protocols*, 7(1) :34–49, jan 2012.
- [MRR<sup>+</sup>53] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6) :1087–1092, dec 1953.
- [MS95] John F Marko and Eric D Siggia. Stretching DNA. *Macromolecules*, 28(26) :8759–8770, dec 1995.
- [MTB<sup>+</sup>10] Manoel Manghi, Catherine Tardin, Julien Baglio, Philippe Rousseau, Laurence Salomé, and Nicolas Destainville. Probing DNA conformational changes with high temporal resolution by tethered particle motion. *Physical biology*, 7(4) :046003, jan 2010.
- [NN08] Keir C Neuman and Attila Nagy. Single-molecule force spectroscopy : optical tweezers, magnetic tweezers and atomic force microscopy. *Nature methods*, 5(6) :491–505, jun 2008.
- [NS76] E Neher and B Sakmann. Noise analysis of drug induced voltage clamp currents in denervated frog muscle fibres. *The Journal of physiology*, 258(3) :705–729, jul 1976.
- [PDT<sup>+</sup>04] Noëlle Pouget, Cynthia Dennis, Catherine Turlan, Mikhail Grigoriev, Michaël Chandler, and Laurence Salomé. Single-particle tracking for DNA tether length monitoring. *Nucleic acids research*, 32(9) :e73, 2004.
- [SFB92] S B Smith, L Finzi, and C Bustamante. Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science (New York, N.Y.)*, 258(5085) :1122–1126, nov 1992.
- [SNPo6] Darren E. Segall, Philip C. Nelson, and Rob Phillips. Volume-exclusion effects in tethered-particle experiments : Bead size matters. *Physical Review Letters*, 96(8) :088306, mar 2006.
- [vSo6] Marian von Smoluchowski. Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Annalen der Physik*, 326(14) :756–780, 1906.
- [WJ57] W W Wood and J D Jacobson. Preliminary Results from a Recalculation of the Monte Carlo Equation of State of Hard Spheres. *The Journal of Chemical Physics*, 27(5) :1207, aug 1957.

- [WP57] W W Wood and F R Parker. Monte Carlo Equation of State of Molecules Interacting with the Lennard-Jones Potential. I. A Supercritical Isotherm at about Twice the Critical Temperature. *The Journal of Chemical Physics*, 27(3) :720, aug 1957.
- [WWPS92] P. C. Weber, J. J. Wendoloski, M. W. Pantoliano, and F. R. Salemme. Crystallographic and thermodynamic comparison of natural and synthetic ligands bound to streptavidin. *Journal of the American Chemical Society*, 114(9) :3197–3200, apr 1992.





## **Troisième partie**

# **Effet Global : variation de la force ionique**



# ÉTUDE EXPÉRIMENTALE ET THÉORIQUE DES VARIATIONS DE LA FORCE IONIQUE SUR L'ADN

## SOMMAIRE

INTRODUCTION . . . . .	143
5.1 LES POLYÉLECTROLYTES . . . . .	145
5.1.1 Électrolyte aqueux . . . . .	145
Théorie de Debye-Hückel . . . . .	145
La longueur de Bjerrum : $l_B$ . . . . .	146
La longueur de Debye : $l_d$ ou $\kappa^{-1}$ . . . . .	146
5.1.2 La force ionique : $I_{ion}$ . . . . .	147
Définition . . . . .	147
Calcul détaillé de la force ionique utilisée durant nos expériences TPM . . . . .	147
Gammes de force ionique utilisées . . . . .	148
Quelle est la variation de $\kappa^{-1}$ associée à notre gamme de force io- nique ? . . . . .	149
5.1.3 Polyélectrolyte, quelques grandeurs caractéristiques . . . . .	150
Distance curviligne entre deux charges . . . . .	150
Densité linéique de charge . . . . .	151
5.1.4 Théorie de Manning ou la condensation des contre-ions . . . . .	151
Paramètre de Manning . . . . .	152
Limites de l'effet de condensation . . . . .	152
Taux de charges neutralisées . . . . .	153
5.2 MESURE DE LA VARIATION DES PROPRIÉTÉS PHYSIQUES DE MOLÉCULES D'ADN EN FONCTION DE LA FORCE IONIQUE . . . . .	154
5.2.1 Variation de la distance bout-à-bout des complexes ADN/particule . . . . .	154
5.2.2 Extraction de la longueur de persistance via les modèles <i>classiques</i> . . . . .	156
5.2.3 Extraction de la longueur de persistance via notre modèle Échan- tillonnage Exact . . . . .	159
Dépendance de $R_{  ,EE}$ en fonction de $L_p$ . . . . .	161
Extraction de la valeur $L_p$ de nos données TPM . . . . .	161
5.3 VARIATION DE $L_p$ . . . . .	163
5.4 VALEUR DE LA LONGUEUR DE PERSISTANCE PRÈS DES CONDITIONS SA- LINES PHYSIOLOGIQUES . . . . .	163

5.5	THÉORIE ODIJK, SKOLNICK, FIXMAN . . . . .	165
5.5.1	Longueur de persistance électrostatique . . . . .	165
5.5.2	Limites de la théorie OSF . . . . .	165
5.5.3	Analyse de nos données par cette théorie . . . . .	166
5.6	PONDÉRATION À FAIBLE $I$ DE LA THÉORIE OSF PAR LE PARAMÈTRE DE MANNING . . . . .	166
5.6.1	Limites de la pondération . . . . .	167
5.6.2	Analyse de nos données par cette théorie . . . . .	167
	Gamme de $Na^+$ . . . . .	168
	Gamme de $Mg^{2+}$ . . . . .	168
5.7	CORRECTION AUX VALEURS DE $I$ INTERMÉDIAIRES PAR UNE THÉORIE DE CHAMP MOYEN NON LINÉAIRE . . . . .	170
5.7.1	Dépendance de $\zeta$ en fonction de $I$ . . . . .	170
5.7.2	Limites de la théorie de champ moyen . . . . .	171
5.7.3	Analyse de nos données par cette théorie . . . . .	171
5.8	PRISE EN COMPTE DES FORCES D'ÉTIREMENT DANS UNE NOUVELLE THÉORIE DE MANNING . . . . .	172
5.8.1	Analyse de nos données par cette théorie . . . . .	173
5.9	PROPOSITION D'UNE FORMULE D'INTERPOLATION . . . . .	175
5.9.1	Analyse de nos données par cette théorie . . . . .	175
5.10	MESURE D'UN EFFET SÉQUENCE PAR L'APPROCHE TPM . . . . .	177
5.10.1	Variation de la distance bout-à-bout des complexes ADN/particule . . . . .	177
5.10.2	Variation de la longueur de persistance . . . . .	178
5.10.3	Ajustement via la nouvelle théorie de Manning . . . . .	179
5.11	OBSERVATION SUR LA VARIATION DU TEMPS DE RELAXATION . . . . .	182
	CONCLUSION . . . . .	183
	ARTICLE . . . . .	185

## INTRODUCTION

**I**N *vivo*, il est établi que la présence d'ions et de gradients de concentration en ions sont importants dans le déroulement de processus biologiques. L'exemple le plus flagrant correspond à la possibilité de **dénaturation chimique** de la double hélice d'ADN. En effet, depuis les années 60, la dépendance de la température de dénaturation de l'ADN en fonction de la concentration saline est étudiée [Sch65]. La force ionique influe sur cette température caractéristique de dénaturation de la double hélice. L'effet de la force ionique sur des propriétés physiques de la molécule d'ADN, telle que la longueur de persistance, est ainsi étudié depuis longtemps. Par contre, la quantification et l'explication de cet effet restent encore aujourd'hui sujettes à controverse.

La molécule d'ADN est sensible aux ions, de par sa nature même. En effet, comme nous l'avons vu, le squelette phosphaté de la molécule d'ADN est chargé négativement. Ceci confère à ce bio-polymère la capacité d'interagir avec des molécules chargées telles que les protéines, mais aussi, et tout simplement, d'interagir avec des ions et contre-ions présents en solution. La vision la plus simple de l'effet induit par ces interactions ioniques, est basée sur l'idée suivante. Si nous considérons une molécule d'ADN isolée, hors de tout milieu, les charges négatives de squelette se repoussent mutuellement sous l'effet des interactions électrostatiques. La traduction de ce phénomène conduit à un étirement de la molécule et donc à une augmentation de la **distance bout-à-bout**. Ceci induit une augmentation de rigidité de l'ADN, autrement dit de sa **longueur de persistance**.

Si maintenant, l'ADN est placé dans une solution chargée incluant, par exemple, des cations  $Na^+$ , ces derniers viennent interagir avec les ions phosphates négatifs de la structure interne du polymère. Ceci crée un phénomène d'**écranage**, diminuant la portée des interactions électrostatiques répulsives entre les anions phosphates. Tout cela induit une modification des conformations accessibles par le polymère, autrement dit une diminution de la **longueur de persistance** par rapport au cas précédent.

Ainsi, depuis la fin des années 70, de nombreuses études expérimentales ont exploré l'effet de variation de la force ionique. La première étude centrée sur cette question fut réalisée en 1978 par Harrington [Har78], via la technique de birefringence par flux qui permet d'extraire la mesure du rayon de giration du polymère. La variation de la longueur de persistance extraite en fonction de la force ionique n'est pas simplement explicable par une loi de puissance, notamment aux concentrations faibles et intermédiaires en ions  $Na^+$ . Depuis, de nombreuses études expérimentales ont suivi afin d'explorer l'influence des ions présents en solution, dans différentes gammes de concentration, sur les propriétés physiques de l'ADN. Aujourd'hui, le bilan de ces études fait apparaître une grande diversité de résultats expérimentaux dont il est difficile d'extraire un comportement général. En effet, ces études mettent en exergue l'existence de différents comportements de l'ADN, principalement en présence de fortes forces ioniques [Sav12].

De plus, les nombreuses études théoriques, que se soit la théorie d'Odjik-Skolnick-Fixman (OSF) [Odi77, FS78] ou encore le modèle de condensation de Manning ne permettent pas de rendre entièrement compte des variations de la longueur de persistance observées [Man69a, Man69b, Man81]. En effet, ces dernières décrivent convenablement le profil des données expérimentales sur des gammes restreintes de force ionique. Nous verrons que le modèle OSF permet d'exploiter les données sur la gamme à forte concentration en sel, tandis que la théorie de Manning décrit de façon plus adaptée les données à faible force ionique.

Cette division de la réponse physique de l'ADN entre ces deux gammes de force ionique soulève toujours de nombreuses questions et ouvre un large champ d'investigations. Ainsi, durant cette dernière décennie, Manning a repris sa théorie afin de rendre compte de la dépendance de  $L_p$  en fonction de la force ionique sur l'ensemble des concentrations [Man06]. La nouvelle théorie qu'il établit en 2006 prend en compte l'étirement électrostatique interne à la molécule due à la répulsion entre les charges portées le long de la chaîne d'ADN. Cette théorie permet de mieux réconcilier les variations des propriétés physiques de l'ADN à fort et à faible sel.

C'est dans ce contexte expérimental et théorique que s'inscrit notre travail. De nombreuses questions restent ouvertes.

La présence de sel en solution peut donc avoir un effet notable sur l'amplitude du mouvement des complexes ADN/particule observée expérimentalement en TPM. Si l'observation d'un tel effet est possible dans nos expériences TPM, il est nécessaire de les quantifier avant d'aller plus en avant dans l'observation d'un effet local. Ainsi il nous sera possible d'identifier de façon séparée l'impact induit par une modification de la force ionique de celui induit par d'autres effets et conditions appliquées à la molécule d'ADN.

Quelles sont les observables qui nous sont accessibles via les expériences TPM? Quels sont donc les effets de la présence d'ions en solution sur les propriétés physiques de l'ADN? Existe-il différents régimes de réponse de l'ADN en fonction de la concentration des ions présents en solution? L'effet est-il semblable si les ions sont monovalents ou divalents? Via l'apport des simulations, nous est-il possible de modéliser cette dépendance? De plus, nous est-il possible de résoudre le problème inverse? Quelles sont les théories existantes sur le sujet? Quelles sont leurs limites? L'ensemble de ces questions seront traitées dans les pages suivantes.

Ce chapitre est développé selon un format spécifique. En effet, un grand nombre des résultats de cette partie sont publiés. Nous avons donc décidé, afin de ne pas simplement traduire la publication, d'ajouter cette dernière à la fin de ce chapitre afin de présenter un travail complet et cohérent. Ainsi, les théories fondamentales et principales seront précisées mais les paramètres expérimentaux, déjà renseignés dans l'article ne seront pas détaillés. Des compléments sont apportés dans ce chapitre sur des points non développés (ou pas entièrement) dans la publication, et inversement, certains points et aspects de la discussion sont entièrement traités dans l'article.

## 5.1 LES POLYÉLECTROLYTES

Avant de rentrer dans le cœur des travaux et connaissances autour des polyélectrolytes, avec notamment la description du paramètre de Manning et de la théorie OSF, définissons un certain nombre de paramètres qui nous seront utiles. Tout d'abord rappelons quelques notions liées aux électrolytes et la théorie de Debye-Hückel.

### 5.1.1 Électrolyte aqueux

Un ion en solution peut-être considéré comme une charge isolée. Afin de respecter la règle d'électronégativité, une répartition homogène de charge de signe contraire entoure l'ion. Si nous considérons l'ion ainsi que ses contre-ions, dans un élément de volume suffisamment grand, ce système est électriquement neutre. Il existe donc des fluctuations de charge et de potentiel électrique important dans ce volume.

#### Théorie de Debye-Hückel

La théorie de Debye-Hückel propose une représentation des ions afin de prendre en compte ce phénomène. En effet, ce modèle considère la distribution des ions, supposée isotrope, autour d'un ion central, pris au hasard. De plus, la densité d'ions associée à cette distribution de charge, à symétrie sphérique, suit la statistique de Boltzmann.

Nous négligerons la taille des ions, ce qui est théoriquement valable si la distance moyenne entre les ions,  $d_{ion-ion}$ , est grande devant la taille de ces derniers, ce qui correspond à se placer à faible concentration en ions.

Différents types de rayons peuvent être associés aux ions. Les rayons atomiques,  $r_a$ , aussi nommés rayons de Bohr car dérivés directement de la théorie de Bohr qui caractérise la taille du nuage électronique de l'atome. Le rayon cristallin,  $r_c$ , ou aussi nommé rayon de Pauling. Ce rayon correspond au rayon ionique du cristal nu. De plus, nous pouvons aussi relever la valeur du rayon de Stokes,  $r_s$ . Ce dernier prend en compte la couche d'hydratation présente autour de l'ion solvaté qui influe lors du déplacement de ce dernier dans un fluide. Dans de l'eau à 25°C, les rayons des ions sur lesquels nous avons travaillé dans ce projet sont répertoriés dans le tableau suivant.

Ion	Rayon atomique : $r_a$ nm	Rayon cristallin : $r_c$ nm	Rayon de Stokes : $r_s$ nm
$Na^+$	0.169	0.095	0.184
$Cl^-$	0.202	0.181	0.121
$Mg^{2+}$	0.142	0.065	0.341

TABLE 5.1 – Valeurs des différents rayons associés aux ions utilisés, extrait de [HA06].

La théorie de Debye-Hückel reste valable même lorsque les ions sont très proches les uns de autres. Notamment l'étude par simulation Monte-Carlo de Olivares et McQuarrie a montré en 1975 que cette théorie reste valable dans le cas où  $d_{ion-ion}$  est de l'ordre de l'angström [OM75].

**La longueur de Bjerrum :  $l_B$** 

**La longueur de Bjerrum** correspond à la longueur pour laquelle deux charges électriques  $e$  dans un solvant de constante diélectrique  $\epsilon$ , interagissent l'une avec l'autre avec une énergie égale à  $k_B T$ .

$$l_B = \frac{e^2}{4\pi\epsilon k_B T} \quad (5.1)$$

- $e = 1.602 \cdot 10^{-19}$ , A.s
- $\epsilon = \epsilon_r \epsilon_0 = 78.3 \times 8.85 \times 10^{-12}$ , A<sup>2</sup>. s<sup>4</sup>. kg<sup>-1</sup>. m<sup>-3</sup>
- $k_B T = 4.2 \times 10^{-21}$ , J.K<sup>-1</sup>s à température ambiante

Autrement dit, si la distance entre deux charges élémentaires est égale à la **longueur de Bjerrum**, alors l'**interaction électrostatique** entre ces deux charges correspond strictement à l'énergie associée à l'**agitation thermique**. À température ambiante, la **longueur de Bjerrum** entre deux charges présentes dans de l'eau est de 0.7135 nm [Odi77].

**La longueur de Debye :  $l_d$  ou  $\kappa^{-1}$** 

**La longueur de Debye** correspond à la distance typique pour laquelle les charges portées par une molécule chargée sont écrantées par les autres charges électriques présentes en solution. Les ions positifs et négatifs en solution forment un nuage autour des charges négatives portées par l'ADN, diminuant ainsi la portée de ces dernières. Ceci induit donc la diminution de la portée des interactions électrostatiques répulsives entre les groupements phosphates.

Le potentiel effectif associé à la théorie de Debye-Hückel correspond au potentiel coulombien existant entre deux charges espacées l'une de l'autre d'une distance  $r$ , avec  $r \gg r_{ion}$ , selon la forme suivante :

$$\frac{e}{k_B T} \Phi(r) = \frac{l_B}{r} \exp(-\kappa r) \quad (5.2)$$

Le potentiel électrostatique induit par ce nuage ainsi formé décroît exponentiellement. La longueur typique de cette décroissance est la longueur de Debye. Cette dernière est une fonction décroissante de la longueur de Bjerrum ainsi que de la force ionique  $I$  que nous définirons en suivant.

$$\kappa_{[m^{-2}]}^2 = 8\pi l_B I_{[ion/m^3]} \quad (5.3)$$

$$\Rightarrow \kappa_{[m]}^{-1} = \frac{1}{\sqrt{8\pi l_B I_{[ion/m^3]}}} \quad (5.4)$$

Ainsi, la **longueur de Debye**,  $\kappa^{-1}$ , longueur sur laquelle les interactions électrostatiques influent, diminue lorsque la salinité du milieu augmente. Dans une solution aqueuse à 25°C et pour un sel monovalent,  $\kappa^{-1}$  peut donc s'écrire [FS78] :

$$\Rightarrow \kappa_{[nm]}^{-1} = \frac{0.304}{\sqrt{I_{[mol/L]}}} \quad (5.5)$$



### 5.1.2 La force ionique : $I_{ion}$

#### Définition

La **force ionique** de la solution permet de caractériser la quantité et l'activité des ions présents en solution. Cette dernière correspond tout simplement à la somme des concentrations des ions présents en solution pondérée par le carré de leur charge respective. Son unité internationale est  $\text{ion}/\text{m}^3$ , comme dans l'Eq. 5.6, mais elle est très souvent exprimée en  $\text{mol}/\text{L}$ , Eq. 5.7.

$$I_{ion, [\text{ion}/\text{m}^3]} = I_{[\text{mol}/\text{L}]} \times N_a \times 10^3 \quad (5.6)$$

$$\Rightarrow = \frac{1}{2} \sum_i Z_i^2 C_i [\text{mol}/\text{L}] \quad (5.7)$$

- $C_i$  la concentration de l'espèce  $i$  en  $\text{mol} \cdot \text{L}^{-1}$
- $Z_i$  la valence de l'espèce  $i$
- $N_a$  nombre d'Avogadro en  $\text{mol}^{-1}$ .

#### Calcul détaillé de la force ionique utilisée durant nos expériences TPM

Dans nos expériences, nous utilisons une solution saline de tampon phosphate, PBS, constituée de  $\text{KH}_2\text{PO}_4$  et de  $\text{Na}_2\text{HPO}_4$  aux concentrations respectives  $a = 1,06 \text{ mM}$  et  $b = 3,00 \text{ mM}$ . En fonction de la nature des ions dont nous cherchons à étudier l'influence, nous ajoutons, soit une concentration  $c$  de sel monovalent  $\text{NaCl}$ , soit une concentration  $d$  de sel divalent  $\text{MgCl}_2$ .

La force ionique mise en jeu dans nos expériences avec des ions monovalents seuls, s'écrit :

$$\begin{aligned} I_{ion, [\text{mol}/\text{L}]}^{Na^+} &= \frac{1}{2} \left( C_{sel}^{[NaCl]} + C_{Tampon}^{[K^+]} + C_{Tampon}^{[H_2PO_4^-]} + 2 \cdot C_{Tampon}^{[Na^+]} + Ion_{Tampon}^{[HPO_4^{2-}]} \right) \\ \Rightarrow I_{ion, [\text{mol}/\text{L}]}^{Na^+} &= \frac{1}{2} \left( [Na^+]_c (1^2) + [Cl^-]_c (1^2) + [K^+]_a (1^2) + [H_2PO_4^-]_a (1^2) \right. \\ &\quad \left. + 2 \cdot [Na^+]_b (1^2) + [HPO_4^{2-}]_b (2^2) \right) \\ \Rightarrow I_{ion, [\text{mol}/\text{L}]}^{Na^+} &= \frac{1}{2} (2 \cdot [Na^+]_c + 2 \cdot [K^+]_a + 6 \cdot [Na^+]_b) \\ \Rightarrow I_{ion, [\text{mol}/\text{L}]}^{Na^+} &= ([Na^+]_c + 10.06 \times 10^{-3}) \quad \text{mol/L} \end{aligned}$$

La force ionique mise en jeu dans nos expériences avec des ions divalents ajoutés, s'écrit :

$$\begin{aligned} I_{ion, [\text{mol}/\text{L}]}^{Mg^{2+}} &= \frac{1}{2} \left( C_{sel}^{[MgCl_2]} + C_{Tampon}^{[K^+]} + C_{Tampon}^{[H_2PO_4^-]} + 2 \cdot C_{Tampon}^{[Na^+]} + C_{Tampon}^{[HPO_4^{2-}]} \right) \\ \Rightarrow I_{ion, [\text{mol}/\text{L}]}^{Mg^{2+}} &= \frac{1}{2} \left( [Mg^{2+}]_d (2^2) + 2 \cdot [Cl^-]_d (1^2) + [K^+]_a (1^2) + [H_2PO_4^-]_a (1^2) \right. \\ &\quad \left. + 2 \cdot [Na^+]_b (1^2) + [HPO_4^{2-}]_b (2^2) \right) \\ \Rightarrow I_{ion, [\text{mol}/\text{L}]}^{Mg^{2+}} &= \frac{1}{2} (6 \cdot [Mg^{2+}]_d + 2 \cdot [K^+]_a + 6 \cdot [Na^+]_b) \\ \Rightarrow I_{ion, [\text{mol}/\text{L}]}^{Mg^{2+}} &= (3 \cdot [Mg^{2+}]_d + 10.06 \times 10^{-3}) \quad \text{mol/L} \end{aligned}$$

Gammes de force ionique utilisées

La valeur de la force ionique lorsqu’aucun sel supplémentaire n’est ajouté en solution est la suivante :

⇒  $I_{ion,[mol/L]} = (a + 3b) = 10.06 \times 10^{-3} [mol/L]$

Cette solution sera notre valeur de référence associée aux conditions *zéro* sel ajouté. Dans la suite, nous désignerons cette solution via le terme *zero-salt-buffer*.

1	$C_{KH_2PO_4}$	en mmol/L	$a$
2	$C_{Na_2HPO_4}$	en mmol/L	$b$
3	Ion		<i>zero-salt-buffer</i>
4	$C_{ion}$ ajoutée	en mmol/L	∅
5	$I_{ion}$	en mmol/L	10.1
6	Inverse de la Force ionique : $1/I_{ion}$	en L/mol	99.4

Les ions  $Na^+$  présents dans ce milieu sont uniquement apportés par la solution tampon utilisée. La composition de ce tampon PBS implique donc une force ionique minimale de 10.1 mmol/L, ou autrement dit la gamme de  $1/I$  prospectée ne peut être supérieure à 99.4 L/mol. Si nous souhaitons sonder des forces ioniques plus faibles que 10.1mmol/L il serait donc nécessaire de modifier la composition de la solution tampon utilisée.

L’idée de réaliser une mesure à une force ionique la plus faible possible semble intéressante. Autrement dit, l’essentiel des ions présents en solution contribuant à la force ionique correspondent aux contre-ions intervenant lors de la solubilisation la molécule d’ADN. Pour cela notre milieu est uniquement constitué d’eau *ultra-pure* et des surfactants habituels. Cette eau distillée est filtrée pour diminuer de façon conséquente les ions en solution. Ainsi, à 25°C une *eau ultra-pure* possède une conductivité de  $0.05 \times 10^{-6}$  S/m, contre une conductivité de  $\sim 500$  S/m pour l’eau potable. Une solution constituée uniquement d’eau distillée ne possède pas un pH bien contrôlé. Néanmoins, nous avons réalisé des mesures dans un tel milieu complété par des polymères non chargés, des surfactants, afin de limiter les collages non spécifiques et irréversibles de nos complexes ADN/particule sur la surface. Malheureusement la réalisation d’une telle mesure, à une force ionique le plus proche de *zéro* possible, ne fut pas un succès. En effet, l’ensemble des complexes ADN/particule se sont collés non spécifiquement sur la surface malgré la présence de surfactants. Ceci souligne bien l’importance *in vivo*, de tous les composants du milieu.

L’effet des ions **monovalent**  $Na^+$  sera testé sur la gamme de concentrations suivantes, les lignes numérotées de 1 à 6 correspondant aux précédentes :

1	$a$													
2	$b$													
3	<i>NaCl</i>													
4	5	10	15	20	25	30	40	50	100	155	500	1500	3000	
5	15.1	20.1	25.1	30.1	35.1	40.1	50.1	60.1	110.1	165.1	510.1	1510.1	3010.1	
6	66.4	49.9	39.99	33.3	28.5	25.0	20.0	16.7	9.1	6.1	2.0	0.7	0.3	

Les concentrations grisées, , correspondent à la plage de concentration en ions monovalent équivalente aux conditions physiologiques.

La solution industrielle servant habituellement de buffer, *Phosphate Buffer Saline Solution de Eurodemex*, durant nos expériences TPM est constituée d'un ajout de 155 mM de  $\text{NaCl}$ . Ceci permet de réaliser une solution tampon proche des conditions salines physiologiques. La valeur de la force ionique associée à  $C_{[\text{NaCl}]} = 155 \text{ mM}$ , sans ajout de  $\text{MgCl}_2$ , est égale à :

$$\Rightarrow I_{\text{ion}, [\text{mol/L}]} = (c + a + 3b) = 165.1 \times 10^{-3} [\text{mol/L}]$$





De même, l'effet des ions **divalents**  $\text{Mg}^{2+}$  sera testé sur la gamme de concentrations suivantes :

1	<i>a</i>											
2	<i>b</i>											
3	<i>MgCl<sub>2</sub></i>											
4	0.15	0.25	0.5	0.75	1	2	5	10	20	50	100	500
5	10.5	10.8	11.6	12.3	13.1	16.1	25.1	40.1	70.1	160.1	310.1	1510.1
6	95.1	92.5	86.5	81.2	76.6	62.3	39.9	25.0	14.3	6.2	3.2	0.7

Valeur de la force ionique pour  $C_{[\text{MgCl}_2]} = 500 \text{ mM}$ , sans ajout de  $\text{NaCl}$

$$\Rightarrow I_{\text{ion}, [\text{mol/L}]} = (3d + a + 3b) = 1510.1 \times 10^{-3} [\text{mol/L}]$$

Dans la suite, lors de la représentation des données, le code couleur suivant sera utilisé :

- ligne pointillée  : données dans la gamme de concentration en sel proche des conditions physiologiques.
- point en  : données à *zero-salt-buffer*.
- point en  : données avec des ions monovalents  $\text{Na}^+$ .
- point en  : données avec des ions divalents  $\text{Mg}^{2+}$ .

Ceci nous permettra de rapidement mettre en évidence la dépendance des caractéristiques physiques de l'ADN en fonction de la concentration en sels ainsi que de la nature des ions présents en solution.

### Quelle est la variation de $\kappa^{-1}$ associée à notre gamme de force ionique ?

La longueur de Bjerrum est indépendante de la force ionique du milieu. La longueur de Debye, quant à elle, est inversement proportionnelle à cette grandeur. Afin de visualiser pleinement la variation de  $\kappa^{-1}$ , Fig. 5.1 et 5.2, sur nos gammes de concentrations en sels, nous ajoutons des symboles pour situer les concentrations étudiées durant cette étude.

La longueur de Debye diminue fortement avec l'augmentation de la concentration en ions présents en solution sur la gamme allant des faibles de force ionique aux  $I$  associés aux conditions physiologiques, autrement dit pour  $I < 0.150 \text{ mol/L}$ . Ensuite, lorsque  $I > 0.150 \text{ mol/L}$ ,  $\kappa^{-1}$  varie plus lentement. Ceci montre bien que

plus il y a d'ions présents en solution, plus l'effet d'écrantage est prépondérant. Ceci induit une répulsion entre les charges de plus en plus faible.

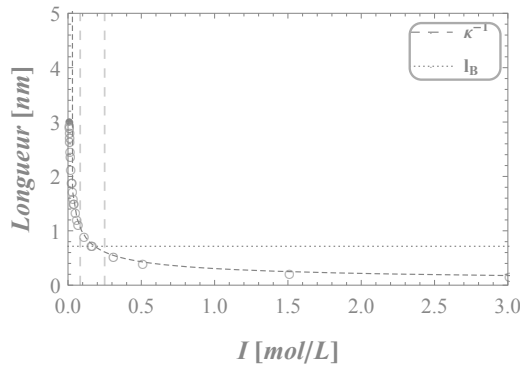


Fig. 5.1 – Variation de  $l_B$  et  $\kappa^{-1}$  en fonction de la force ionique.

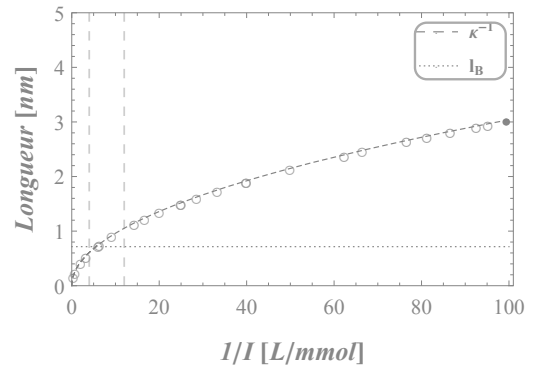


Fig. 5.2 – Variation de  $l_B$  et  $\kappa^{-1}$  en fonction de l'inverse de la force ionique.

La transposition de ce phénomène physique sur notre objet d'étude qu'est l'ADN est la suivante. La répulsion entre les charges négatives portées par le squelette de l'ADN est de plus en plus faible lorsque la force ionique augmente. Autrement dit, la force de répulsion entre ces charges négatives diminue, réduisant ainsi la rigidité de l'ADN. La longueur bout-à-bout effective de celui-ci,  $\langle R^2 \rangle$ , s'en voit aussi diminuée, distance statistique moyenne entre les extrémités de l'ADN à ne pas confondre avec la longueur de contour de la molécule,  $L$ . Cette diminution doit donc être observable via nos mesures expérimentales de molécules uniques.

À titre indicatif, la longueur de Bjerrum observe une diminution relative de 18% lorsque la température diminue de 25°C à 80°C. Ceci induit une augmentation de la longueur de Debye. Ceci est cohérent avec le gain en entropie de translation des ions. En effet, une augmentation de l'agitation thermique rendant compte de l'augmentation du mouvement brownien des particules, tend à limiter le rapprochement des ions les uns des autres. Cette augmentation de température induit une augmentation relative de la valeur de  $\kappa^{-1}$  de 14%.

### 5.1.3 Polyélectrolyte, quelques grandeurs caractéristiques

Un **polyélectrolyte** est un polymère qui dans une solution polaire, autrement dit composée d'électrolytes, présente des charges le long de la chaîne.

Soit un polymère de longueur de contour  $L = Na$ , toujours constitué de  $N$  monomère de taille  $a$ .

#### Distance curviligne entre deux charges

Soit,  $A$  la distance curviligne moyenne entre deux charges portées par la chaîne.

Afin d'illustrer les paramètres associés à un polyélectrolyte nous reprenons notre représentation d'un polymère sous sa forme de pelote aléatoire. Les charges négatives portées par la molécule, représentées en rouge sur Fig. 5.3, sont espacées de façon homogène le long de la chaîne.

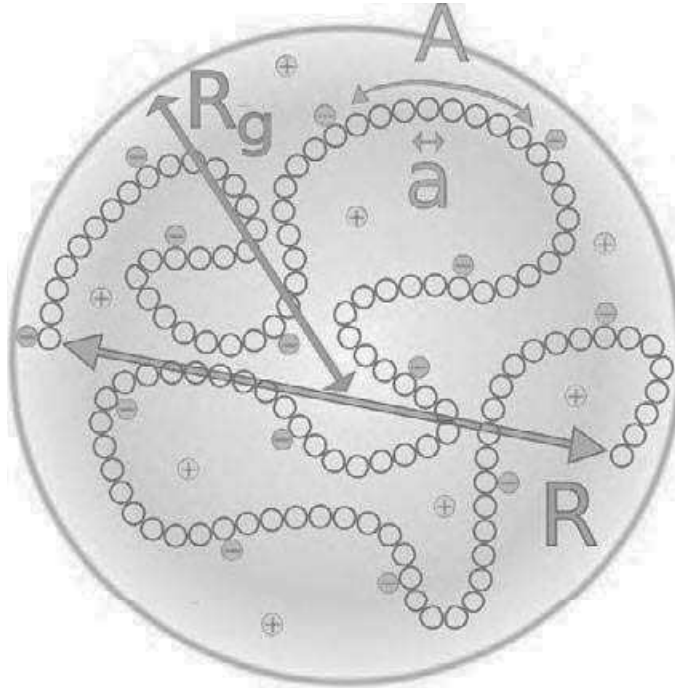


Fig. 5.3 – Représentation d'un polyélectrolyte et de ses longueurs caractéristiques

Intéressons-nous aux caractéristiques typiques d'un ADN de structure hélicoïdale B. Sous la forme de double hélice, cet **ADN nu en solution**, possède une charge négative portée par le groupement phosphate de chaque base azotée. C'est à dire deux électrons par paire de base, ou autrement dit une charge négative espacée toutes les demi-paires de bases. Ainsi la distance moyenne entre deux charges portées par cet ADN,  $A_{ds}$ , est égale à :

$$A_{ds} = \frac{0.34}{2} = 0.17 \text{ nm} \quad (5.8)$$

#### Densité linéique de charge

L'ADN peut-être modélisé localement par un cylindre de rayon  $R_{ADN} = 1 \text{ nm}$ , dont les charges négatives sont uniformément distribuées sur sa surface. La **distance curviligne entre 2 charges portées par la chaîne** nous permet de définir la **densité linéique de charge**  $\rho_0$  par :

$$\rho_0 = \frac{e}{A} \quad (5.9)$$

$$\Rightarrow \rho_{0,A_{ds}} = 0.94 \text{ nC.m}^{-1} \quad (5.10)$$

#### 5.1.4 Théorie de Manning ou la condensation des contre-ions

À la fin des années 60, Manning développe la théorie de **condensation de contre-ions** avant de l'enrichir dans plusieurs papiers successifs [Man69a, Man81, Man69b]. Son travail s'inscrit dans la suite des études des propriétés d'équilibre des électrolytes. La **théorie de Manning** ou **condensation de contre-ions** permet de décrire la neutralisation éventuelle des charges portées le long d'un polyélectrolyte cylindrique due à la présence de contre-ions en solution. Initialement cette théorie est élaborée dans la limite où  $\kappa R_{polyelectrolyte} \rightarrow 0$

Dans le cas où il y a phénomène de condensation des contre-ions le long du polyélectrolyte, le principe associé se comprend aisément. En fonction de la nature du solvant, les charges du polyélectrolyte sont plus ou moins entourées d'un nuage de contre-ions, condensés à une certaine distance du polymère. La taille de ce nuage résulte de l'équilibre entre l'énergie électrostatique, favorisant l'effondrement du nuage sur la chaîne de charges opposées, et l'entropie qui favorise son expansion. La portée de cet effet est dépendante de la longueur d'écrantage des charges, la **longueur de Debye**. Le reste des contre-ions non condensés se trouvent éloignés du polyélectrolyte et sont considérés comme libres.

### Paramètre de Manning

La théorie de Manning tente de quantifier ce phénomène de condensation des contre-ions proches de la surface du polyélectrolyte. Dans cette hypothèse les contre-ions condensés conservent une certaine mobilité le long de la chaîne du polyélectrolyte, qui dans notre cas correspond au polymère d'ADN.

Le polymère est donc considéré comme un cylindre semi-rigide chargé uniformément dans un solvant. Le **paramètre de Manning**,  $\zeta$ , représente la densité de charges portée par le polymère en prenant en compte la constante diélectrique du milieu ainsi que la température, contrairement à l'expression simple de la densité de charge linéique exprimée plus tôt. Le paramètre adimensionné  $\zeta$ , s'exprime tel que :

$$\zeta = \frac{|Z| e^2}{\epsilon k_B T A} \quad (5.11)$$

$$= \frac{|Z| l_B}{A} \quad (5.12)$$

Il est intéressant de noter que le paramètre de Manning dépend directement de la nature du polyélectrolyte via le paramètre  $A$ , ainsi que du solvant via la valence  $Z$  de l'espèce ionique présente. Rappelons que ce dernier ne dépend pas de la concentration en ions car il est calculé dans la limite où  $\kappa R_{polyelectrolyte} \rightarrow 0$ .

### Limites de l'effet de condensation

D'après Manning, ce phénomène de condensation apparaît lorsque la densité de charge dépasse un certain seuil. Ceci permet de définir les bornes suivantes sur la densité de charge afin de rendre compte du phénomène de condensation :

**Si  $\zeta < \frac{1}{|Z|}$  alors il n'y pas de condensation**

La distance entre les charges portées par le polyélectrolyte est supérieure à la longueur de Bjerrum. Les interactions des ions avec le polyélectrolyte peuvent être définies selon l'approximation de **Debye-Hückel** en régime dilué. Autrement dit, l'énergie électrostatique est donnée par le potentiel de **Debye-Hückel**, Eq. 5.2, qui est proportionnel à la longueur d'écrantage de Debye.

Si  $\zeta > \frac{1}{|Z|}$  alors il y a condensation

Dans ce cas, la distance entre les charges portées par le polyélectrolyte est inférieure à la longueur de Bjerrum. L'interaction électrostatique entre les charges est plus grande que l'énergie associée à l'agitation thermique. Le phénomène de condensation des contre-ions apparaît. L'interprétation physique donnée par Manning de ce phénomène, est que le système est instable. Alors, une partie des contre-ions présents en solution se condense autour du polyélectrolyte afin de diminuer la densité de charges effective de ce dernier. En effet, la distance entre les charges portées le long de l'ADN correspond à une distance effective  $A_{eff}$  qui augmente en fonction de l'effet de condensation jusqu'à atteindre sa **valeur seuil**  $A_{eff} = l_B$ . La longueur de Bjerrum apparaît comme la distance minimale possible entre charges.

Le phénomène de neutralisation est d'autant plus important que la valence  $Z$  des contre-ions augmente. D'après Manning, s'ils sont présents en solution, les ions divalents se condensent toujours préférentiellement aux monovalents quelle que soit la concentration de ces derniers.

Le paramètre de Manning associé à un ADN double brin présent dans un solvant ne contenant que des ions monovalent est égal à 4.2, il passe à 8.4 en présence d'ions divalents uniquement. Dans le cas d'un double brins d'ADN,  $\zeta$  est toujours supérieure à  $1/|Z|$ , le phénomène de condensation des contre-ions s'applique.

### Taux de charges neutralisées

Lors du phénomène de condensation, le **taux de charges neutralisées**,  $r$ , correspond au pourcentage de groupements ou charges portées par le polyélectrolyte,  $\alpha_M$ , écrantés dû à l'effet de condensation.

$$r = 1 - \alpha_M = 1 - \frac{1}{\zeta} \quad (5.13)$$

Là encore, appliquons directement cette théorie à notre polymère d'intérêt en présence d'ions monovalent uniquement. Le pourcentage de charges neutralisées  $r$  d'un tel ADN double brin, est de 76% en début de condensation. Autrement dit, le nuage de contre-ions neutralise 76% de la charge négative du polymère. Ce pourcentage, et donc la concentration de contre-ions au voisinage de l'ADN, est notablement important. Comme nous allons le voir, ce phénomène peut être décrit par la compétition entre l'attraction électrostatique et l'agitation thermique, via la théorie de Debye-Hückel qui est une approximation de champ moyen linéarisée.

Cette théorie de Manning est communément étendue afin de décrire le phénomène de condensation lorsque  $\kappa R_{polyelectrolite} \rightarrow 0$ , tout en restant dans la limite des faibles forces ioniques comme nous le verrons plus tard. Ainsi, le reste des co-ions et contre-ions en solution, provenant en partie des concentrations en ions associé à la solution tampon, peut être décrit selon cette approximation de Debye-Hückel.

L'idée principale induite par ce phénomène de condensation réduit donc la charge effective du polymère et influe directement sur ses propriétés physiques et donc les conformations accessibles à la chaîne. Ainsi, un tel phénomène devrait modifier la distance bout-à-bout de l'ADN et donc être observable par des mesures expérimentales utilisant la technique de molécule unique TPM.

## 5.2 MESURE DE LA VARIATION DES PROPRIÉTÉS PHYSIQUES DE MOLÉCULES D'ADN EN FONCTION DE LA FORCE IONIQUE

Afin d'étudier ce phénomène nous avons utilisé différentes longueurs d'ADN. Dans ce travail nous nous sommes concentrés en priorité sur les ADN de longueur  $L_{ADN} = 2060$  et  $1201$  bp, résultats qui ont données lieu à la publication proposée en fin de partie.

Néanmoins, nous avons aussi obtenu des résultats intéressants sur les polymères de taille  $L_{ADN} = 575, 583, 639$  et  $3259$  bp qui sont regroupés dans cette partie. Par contre, ces derniers résultats présentent une erreur de mesure supérieure aux deux premières. En effet, ils correspondent, d'une part, à moins de jours d'expériences et donc un moins grand nombre de trajectoires. D'autre part certaines valeurs de la gamme de force ionique indiquée plus haut, n'ont pas fait l'objet de mesure.

### 5.2.1 Variation de la distance bout-à-bout des complexes ADN/particule

Tout d'abord intéressons-nous à la distance bout-à-bout des complexes ADN/particule, grandeur extraite directement de nos mesures TPM et corrigée de l'effet lié au temps de relaxation des complexes ADN/particule.

Existe-t-il une dépendance de la distance bout-à-bout des complexes ADN/particule, observable et mesurable par TPM, Fig. 3.7.9, en fonction de la force ionique ?

L'ensemble des résultats expérimentaux met en exergue une telle dépendance et ce, quelle que soit la taille de la molécule d'ADN, Fig. 5.4.

Lorsque la force ionique augmente,  $R_{Exp||}$  diminue. En effet, à faible force ionique, les charges négatives portées par l'ADN sont peu écrantées ce qui implique une énergie d'interaction répulsive forte. Cette répulsion tend à étirer le polymère. Lorsque la force ionique augmente, l'effet d'écrantage est plus important. Une image simplifiée correspondrait à l'idée que les charges sur le squelette de l'ADN se voient moins, diminuant ainsi la force de répulsion existant entre-elles. Des sous-parties de la chaîne peuvent alors être dans une conformation plus proches l'une de l'autre, contrairement au cas précédent à fort  $I$ . L'espace des configurations de la chaîne de polymère est moins contraint. Il découle de tout cela que l'ADN est moins étiré.



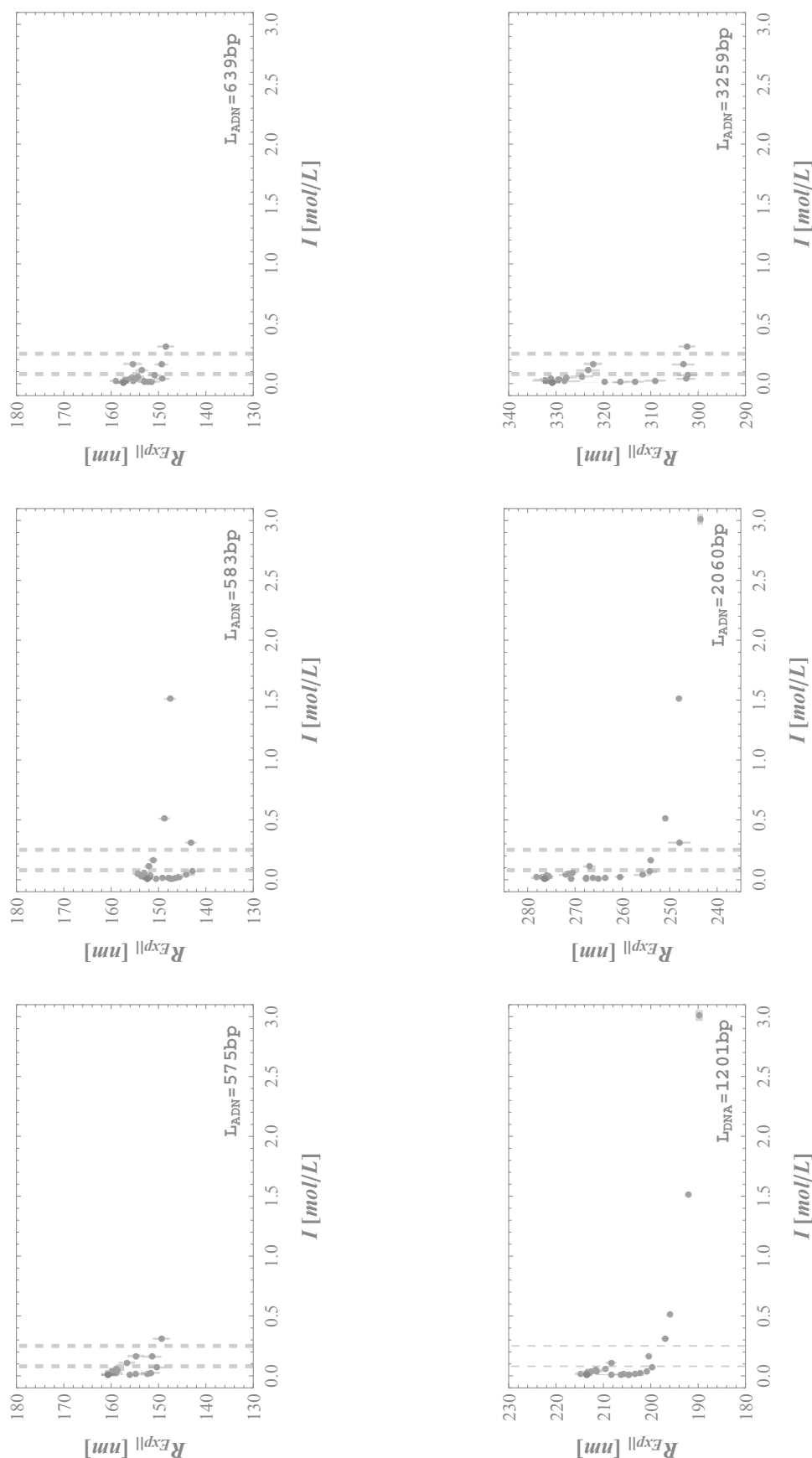


Fig. 5.4 – Variation de la distance bout-à-bout des complexes ADN/particule en fonction de la force ionique pour six longueurs d'ADN. Pour rappel, les symboles rouges correspondent aux mesures en présence d'ions monovalents  $Na^+$ , les bleus aux mesures en présence d'ions divalents  $Mg^{2+}$  tandis que le noir correspond au Zero-salt-buffer. Les barres d'erreurs, si non visibles, sont contenue dans la taille du symboles. Ces dernières sont calculées par estimation bootstrap, comme expliqué dans l'article,  $2\sigma_B$  sont représentés. Dans un souci d'alléger les légendes  $R_{Exp||}$  correspond à la grandeur  $\langle R_{||, (i,i)} \rangle$  définie en 3.7.9

L'amplitude de ce phénomène est plus important pour les ADN de grande taille, comme le choix de représentation de ces données nous aide à le visualiser. Ceci fait écho aux propriétés du volume exclu, qui souligne que toute macromolécule de grande taille est plus sensible, que celle de petite taille, aux effets de volume d'exclusion existant le long de la chaîne même de l'ADN.

De plus, nous pouvons constater que la variation de la grandeur  $R_{Exp||}$  est dépendante de la nature des ions présents en solution. En effet, pour les ions divalents, la longueur bout-à-bout de nos complexes ADN/particule chute plus rapidement en fonction de l'augmentation de  $I$ . À même force ionique, la longueur apparente de l'ADN est plus faible en présence d'ions divalents  $Mg^{2+}$  qu'en présence d'ions monovalents  $Na^+$ . Ceci peut découler de l'augmentation de la valence de l'ion en présence, ou encore de l'augmentation de la taille de ce dernier, cette dernière hypothèse n'a pas été explorée ici.

Pour approfondir notre analyse nous nous focalisons maintenant sur la grandeur physique qu'est la longueur de persistance qui permet de caractériser les propriétés et le comportement de la molécule d'ADN. En effet, le phénomène qui nous intéresse est pleinement pris en compte par la caractérisation de la dépendance de la longueur de persistance en fonction de la force ionique. Afin d'étudier la variation de  $L_p$ , nous avons testé différentes façons d'extraire cette dernière.

### 5.2.2 Extraction de la longueur de persistance via les modèles *classiques*

Tout d'abord, nous avons extrait la contribution de la particule via la méthode classique présentée en Partie 3.8.2, afin d'obtenir la variation de la longueur bout-à-bout de la molécule d'ADN seulement, Fig. 5.5. Ensuite nous utilisons le modèle théorique **WLC**, Eq. 2.22 et 3.8.3, qui permet de décrire la physique du polymère.

Les longueurs de persistance extraites grâce au modèle **WLC**,  $L_{p,WLC}$ , présentent une variation très disparate et éloignée de la valeur attendue  $L_{p,Ref} = 50$  nm, dans la condition saline physiologique, Fig. 5.6. L'utilisation de la méthode présentée par Segall en 2006 pour extraire la contribution de la bille sur les ADN courts par rapport à la taille de la bille, Partie 3.8.2, ne suffit pas à expliquer l'étalement des  $L_{p,WLC}$ . En effet, même pour l'ADN de 2060 bp et de 3259 bp, la longueur de persistance extraite est très grande.

Tout ceci rejoint notre première observation lorsque que nous avons voulu en partant de  $R_{EE||}$  remonter à la longueur de persistance. Cette dernière s'éloigne fortement de la longueur de persistance imposée dans la simulation. Nous nous proposons donc d'utiliser nos simulations par échantillonnage exact afin de résoudre le **problème inverse**.

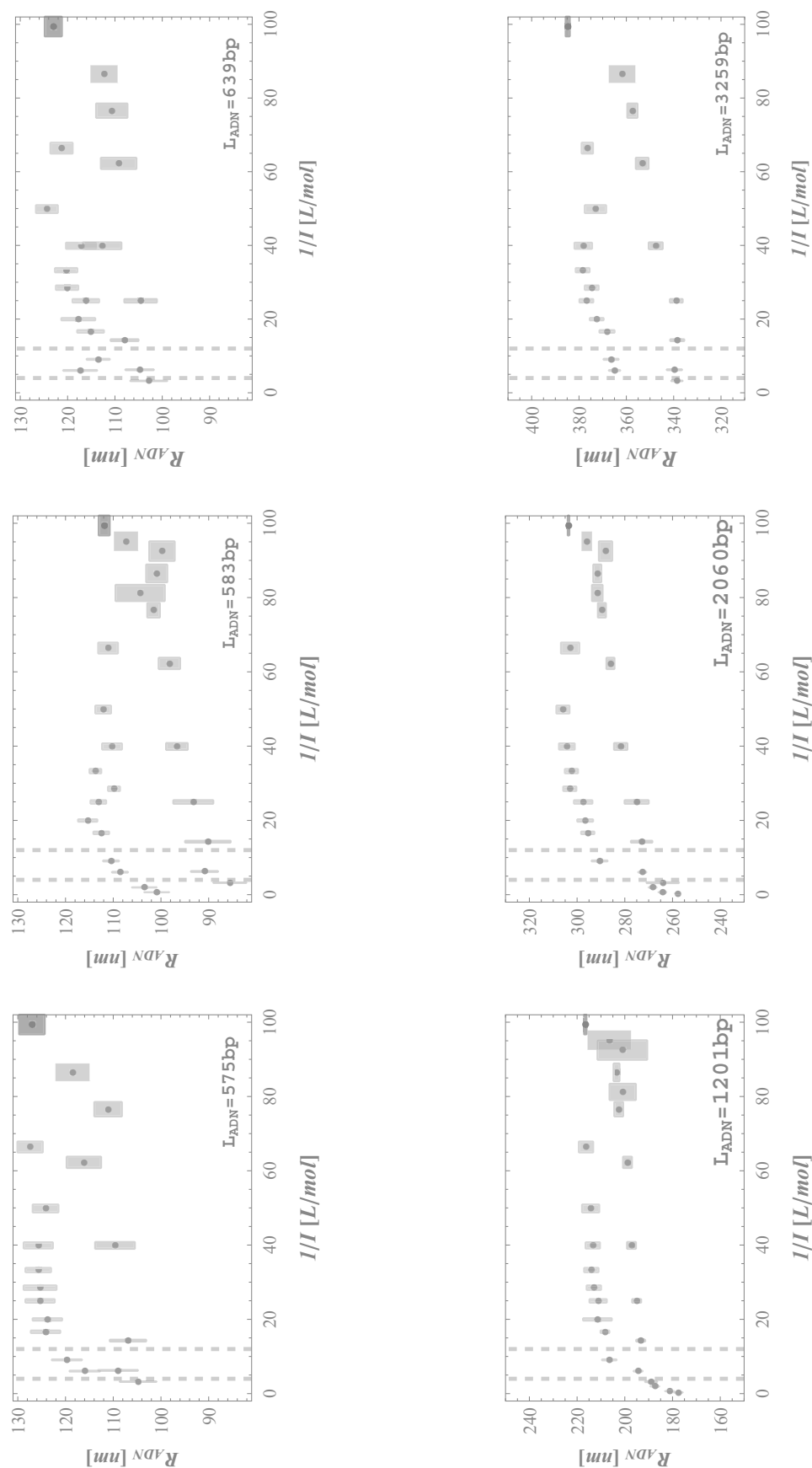


Fig. 5-5 – Variation de la distance bout-à-bout des ADN seuls, en fonction de la force ionique pour six longueurs d'ADN. Les symboles rouges correspondent aux mesures en présence d'ions monovalents  $\text{Na}^+$ , bleus aux mesures en présence d'ions divalents  $\text{Mg}^{2+}$  et noir correspond au Zero-salt-buffer. Les rectangles-barres d'erreur correspondent à la combinaisons de l'erreur de mesure et d'extraction de  $R_{ADN}$  avec l'erreur sur  $I$  découlant de notre précision lors des préparations des solutions tampons, comme décrit dans l'article. Ces derniers sont calculés par estimation bootstrap, comme expliqué dans l'article, si non visible, les rectangles-barres d'erreur sont contenus dans la taille du symbole. L'extraction de  $R_{ADN}$  est définie en 3-7.9

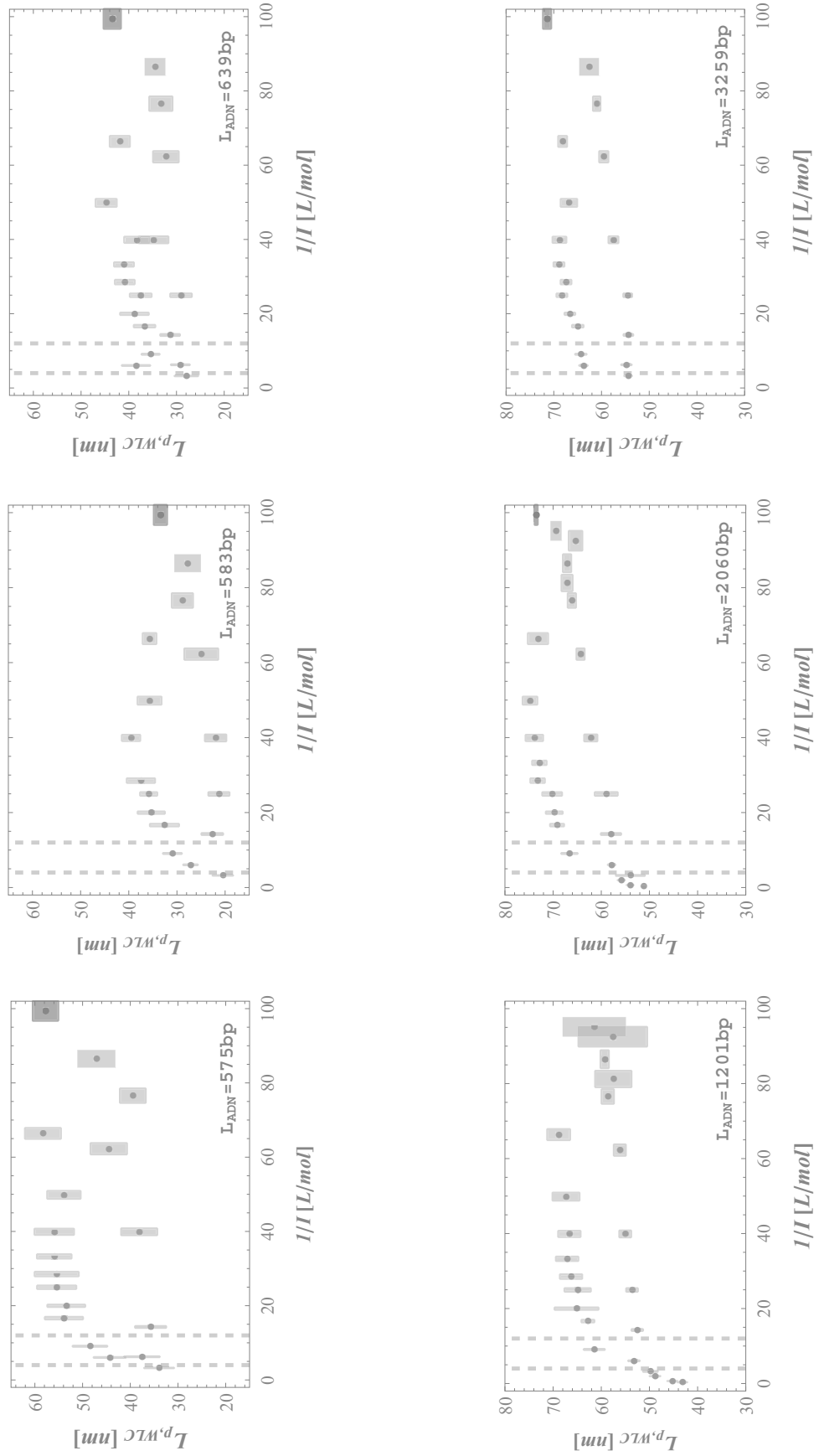


Fig. 5.6 – Variation de la longueur de persistance des ADN en fonction de la force ionique pour six longueurs d'ADN extraite via l'utilisation du modèle WLC 2.22 et 3.8.3.

### 5.2.3 Extraction de la longueur de persistance via notre modèle Échantillonnage Exact

Une variation de la force ionique induit une variation de la longueur de persistance qui se traduit par une variation de la distance bout-à-bout de nos complexes ADN/particule. Nous nous proposons de construire la dépendance de la variation de la distance bout-à-bout de nos complexes simulés ADN/particule en fonction de la variation de la longueur de persistance. Ainsi, pour chaque taille d'ADN utilisée, nous pourrions traduire la variation de la distance bout-à-bout de nos complexes ADN/particule mesurée expérimentalement en terme de variation de  $L_p$ . En d'autres termes, le modèle de simulation EE, nous permet de résoudre le problème inverse et ainsi de remonter à la longueur de persistance sans l'erreur systématique induite par les méthodes d'extraction plus classiques.

Pour cela nous réalisons des simulations par échantillonnage exact modélisant notre système expérimental. Nous simulons les tailles d'ADN suivantes  $L_{ADN} = 575, 583, 639, 1201, 2060$  et  $3259$  bp en imposant la taille des monomères afin de respecter  $a = 6$  bp. Le nombre de monomères pour chaque longueur d'ADN est donc respectivement  $N = (49 + 1), (53 + 1), (100 + 1), (172 + 1), (272 + 1)$ . La gamme de longueurs de persistance prospectée est la suivante  $L_p \in [107 : 207]$  bp ce qui correspond à  $L_p \in [30 : 70]$  nm.

Rappelons que nous ne prenons pas en compte explicitement le solvant dans nos simulations. Les effets liés à la variation du volume exclu en fonction de la force ionique sont donc contenus dans la variation de la longueur de persistance du système. Le volume d'exclusion de chacun des monomères constituant l'ADN ne varie pas en fonction de la longueur de persistance, seul  $L_p$  varie.

Dans nos simulations, nous nous focalisons sur les paramètres de volume d'exclusion suivants :

- Aucun volume exclu : ni mur, ni la chaîne avec elle-même, ni la chaîne avec la particule (Simulation indexée [000], symboles : ▲).
- Pas de volume exclu de la chaîne de polymère avec elle-même (Simulation indexée [101], symboles : ○).
- Tous les volumes exclus sont pris en compte afin de modéliser correctement nos expériences (Simulation indexée [111], symboles : ●).

De plus, nous ajoutons sur le même graphique, pour chaque longueur d'ADN, les expressions analytiques suivantes :

- L'expression du modèle **WLC Discret** associé (Eq. 4.29, courbe pleine)
- L'expression du modèle **WLC** associé (Eq. 2.22 et 3.8.3, courbe hachurée).

La tendance générale des graphiques, Fig. 5.7 est commentée dans l'article traitant des ADN  $L_{ADN} = 1201$  et  $2060$  bp. Par ailleurs, les simulations sur les ADN courts, mettent en exergue les effets de volume exclu induit par la présence du mur, 3.3. La force effective d'allongement vertical exercée par le mur diminue l'espace des configurations accessibles au complexe ADN/particule, diminuant la mesure de sa longueur apparente.

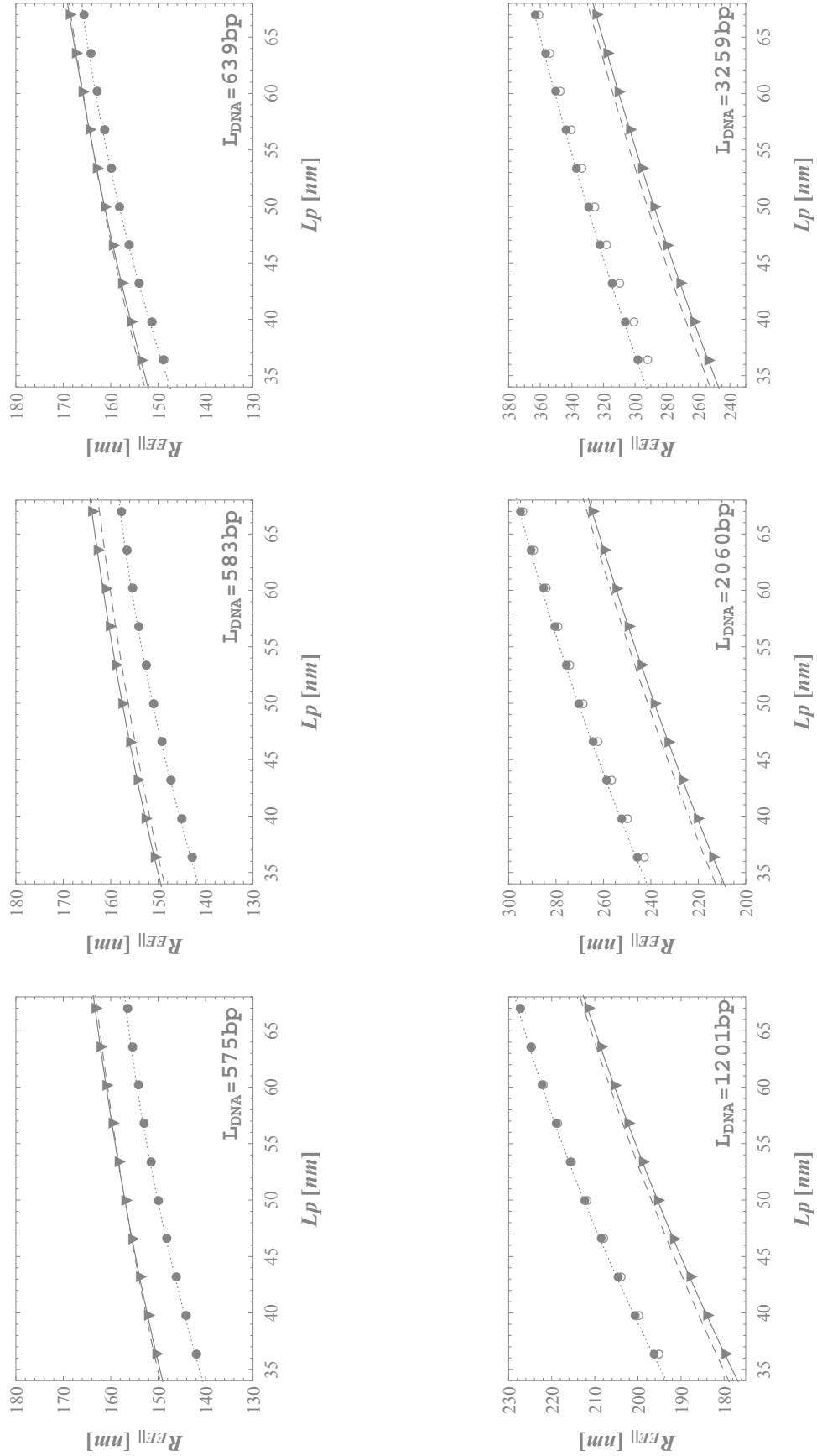


Fig. 5.7 – Données des simulations par échantillonnage exact. Variation de la distance bout-à-bout des systèmes,  $R_{EE||}$ , en fonction de la longueur de persistance,  $L_p$ , imposée dans les simulations, pour six longueurs d'ADN. Symboles  $\blacktriangle$ , simulations sans aucun volume exclu : ni mur, ni la chaîne avec elle-même, ni la chaîne avec la particule. Symboles  $\circ$ , simulations avec volumes exclus sauf de la chaîne de polymère avec elle-même. Symboles  $\bullet$ , simulation avec l'ensemble des volumes exclus. Pour les ADN de petites tailles ces deux dernières série de simulation,  $\circ$  et  $\bullet$ , sont confondues. Courbe continue modèle WLC Discret, courbe en tirets modèle WLC. Courbe pointillées ajustement des données  $\bullet$  via l'expression Eq. 5.14

**Dépendance de  $R_{||,EE}$  en fonction de  $L_p$** 

Les données des simulations imitant notre système expérimental, c'est-à-dire où tous les volumes exclus existant dans le système sont pris en compte, sont ajustées selon la fonction suivante :

$$R_{||,EE} = AL_p^2 + BL_p + C \quad (5.14)$$

L'ajustement de cette dernière est représentée sur la Fig. 5.7, *courbes en pointillées*. Les valeurs des paramètres associés sont résumées dans le tableau ci-dessous.

$Cte$	$L_{ADN} = 575$ nm	$L_{ADN} = 583$ nm	$L_{ADN} = 639$ nm	$L_{ADN} = 1201$ nm	$L_{ADN} = 2060$ nm	$L_{ADN} = 3259$ nm
A	$-6.03 \times 10^{-3}$	$-6.00 \times 10^{-3}$	$-6.83 \times 10^{-3}$	$-9.02 \times 10^{-3}$	$-10.70 \times 10^{-3}$	$-11.52 \times 10^{-3}$
B	1.10	1.11	1.25	1.96	2.71	3.30
C	110.3	110.9	113.0	137.3	162.1	193.9

Une approche complémentaire possible, consisterait à interpoler chacun de ces paramètres en fonction de  $L_{ADN}$ , comme proposé dans l'article [SNP06] afin d'avoir une formule d'interpolation sur  $L_p$  et  $L_{ADN}$ .

**Extraction de la valeur  $L_p$  de nos données TPM**

Nous extrayons la variation de la longueur de persistance grâce à Eq. 5.14.

Au vu du profil de comportement obtenu pour  $L_{ADN} = 575$  bp, Fig. 5.8, nous nous intéresserons à l'étude de ce cas dans la suite. En effet, les longueurs de persistance obtenues s'éloignent des autres données, faisant écho au comportement singulier déjà observé sur  $R_{Exp||}$ , qui est plus grand pour  $L_{ADN} = 575$  bp que pour  $L_{ADN} = 583$  bp, ainsi que lors de l'extraction de  $R_{ADN}$  et  $L_{p,WLC}$ . Pour, l'instant nous laisserons donc de côté cet ADN. Nous reviendrons sur ce dernier dans la partie 5.10.

L'extraction des données est cohérente avec les propriétés physiques attendues pour un polymère d'ADN. En effet, pour l'ensemble des tailles d'ADN étudiées la longueur de persistance se rapproche de  $l_{p,Ref}$  dans la gamme des conditions physiologiques.

De plus, les variations observées ne semblent pas trop dépendantes de la séquence en opposition aux valeurs de  $L_{p,WLC}$  obtenues précédemment. L'utilisation de la méthode d'extraction précédente semblait associer des valeurs de  $L_p$  plus faibles aux ADN de petites tailles, tandis que des plus grandes valeurs de  $L_p$  étaient obtenues pour les longues molécules d'ADN.

Tout ceci renforce la confiance dans la méthode d'extraction basée sur la résolution du problème inverse via des simulations numériques d'un modèle d'échantillonnage exact.

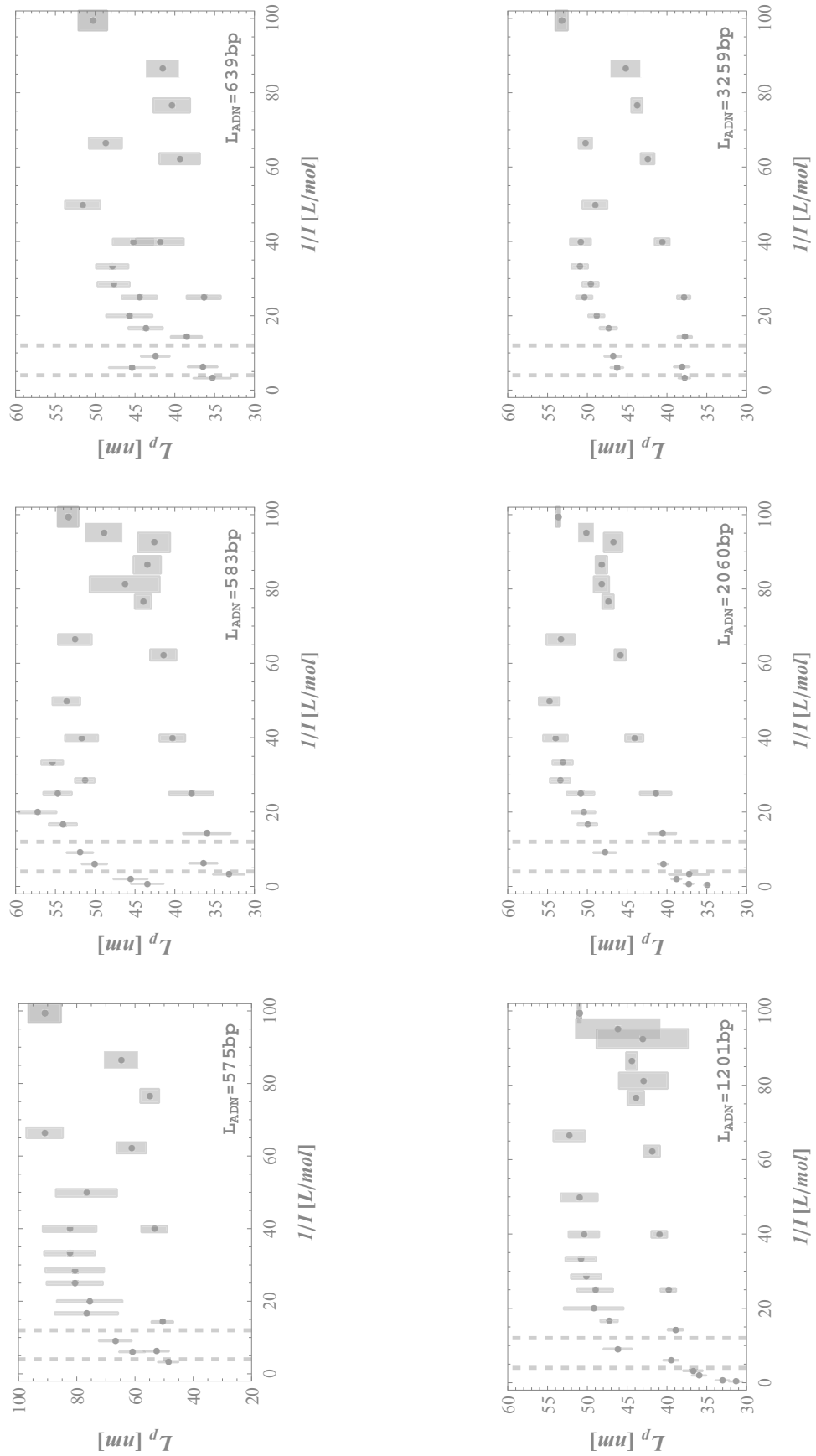


Fig. 5.8 – Variation de la longueur de persistance des ADN en fonction de la force ionique pour six longueurs d'ADN.



### 5.3 VARIATION DE $L_p$

La superposition de l'ensemble de nos données nous permet de mieux juger de l'étalement de la variation de  $L_p$  obtenue, Fig. 5.9.

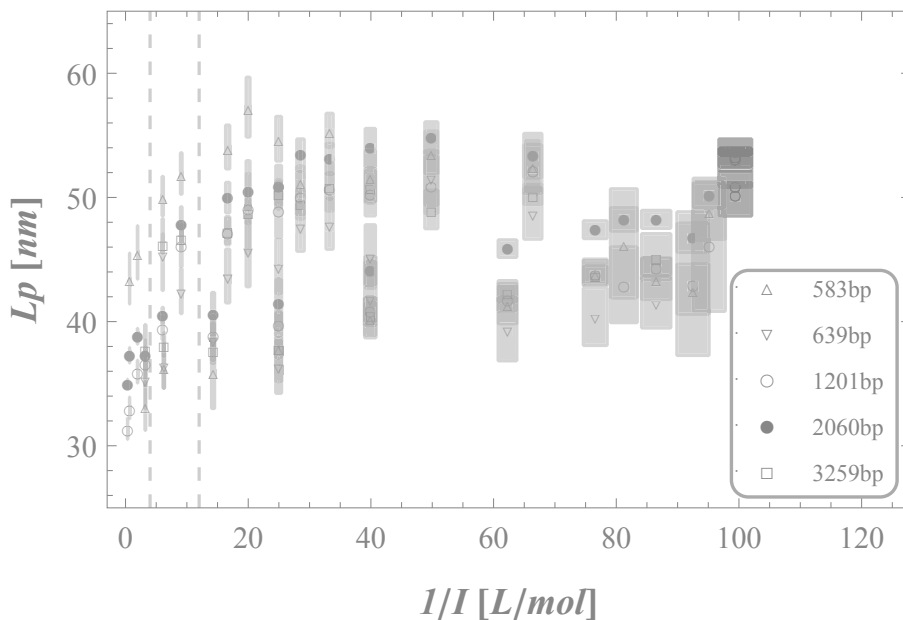


Fig. 5.9 – Variation de  $L_p$  obtenue pour  $L_{ADN} = 583, 639, 1201, 2060$  et  $3259$  bp sur notre gamme de force ionique induite par des ions  $Na^+$ , symboles rouges, et  $Mg^{2+}$ , symboles bleus, les symboles noirs correspondent aux mesures à Zero-salt-buffer)

La dépendance de  $L_p$  est différente en fonction de la nature des ions. Lorsque que  $I$  augmente, autrement dit lorsque  $1/I$  diminue, à cause de la présence d'ions monovalents  $Na^+$ , la longueur de persistance diminue progressivement.  $L_p$  chute plus fortement et brusquement à fort  $I$ , autrement dit à faible  $1/I$ . La variation de la longueur de persistance de l'ADN n'est pas saturée lors de l'exploration des fortes forces ioniques.

Ce profil n'est pas identique lorsque l'ADN est en présence d'ions divalents.  $L_p$  chute plus brusquement dès que des ions divalents sont ajoutés en solution. En effet, nous pouvons observer une forte chute de la longueur de persistance entre la valeur obtenue dans les conditions dites Zero-Salt et la première mesure avec  $Mg^{2+}$ .

Avant d'explorer les théories associées à ce phénomène, nous comparons nos résultats avec les données de la littérature.

### 5.4 VALEUR DE LA LONGUEUR DE PERSISTANCE PRÈS DES CONDITIONS SALINES PHYSIOLOGIQUES

Le graphique suivant, Fig. 5.10, regroupe l'ensemble des données existant dans la littérature sur des gammes de forces ioniques induites par la présence d'ions  $Na^+$ , sans présence d'ions divalents, comparables aux nôtres (se référer à la publication pour plus de précisions).

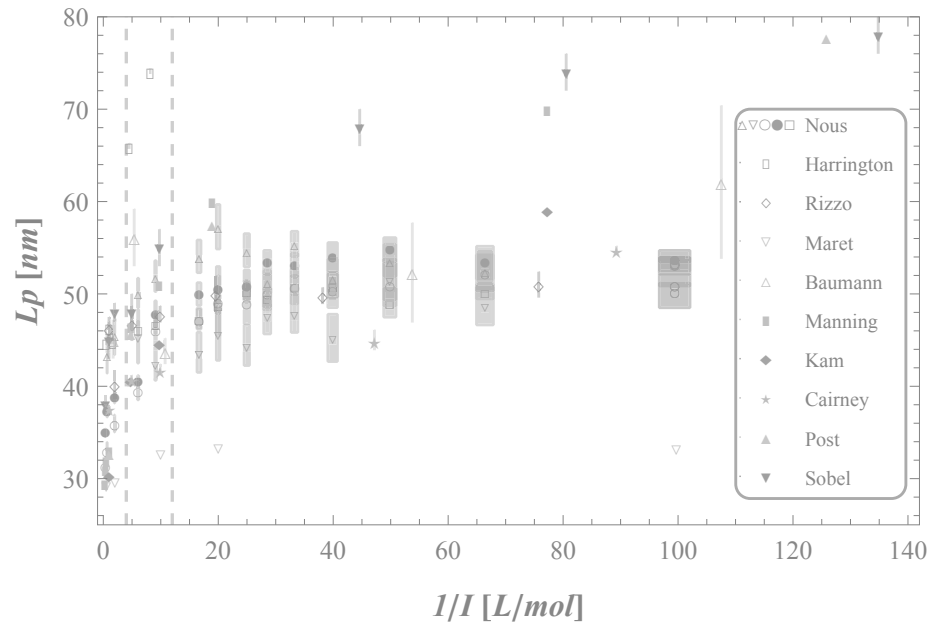


Fig. 5.10 – Comparaison de la variation de  $L_p$  obtenue pour  $L_{ADN} = 583, 639, 1201, 2060$  et  $3259$  bp sur notre gamme de force ionique induite par des ions  $\text{Na}^+$  avec les données de la littérature.

La variation de  $L_p$  précédemment obtenue dans la littérature sur la gamme de  $I$  appartenant seulement à  $[\sim 85 : 200]$  mmol/L, autrement dit pour  $1/I \in [4 : 12]$  L/mmol est de l'ordre de  $48 \pm 6$  nm. Cette valeur est en bon accord avec notre mesure moyenne, toutes tailles de longueur d'ADN confondues, de  $46.6 \pm 4$  nm.

Les quelques données existantes dans la littérature sur des gammes de force ionique induite par la présence d'ion  $\text{Mg}^{2+}$ , comparables aux nôtres, sont regroupées sur la Fig. 5.11.

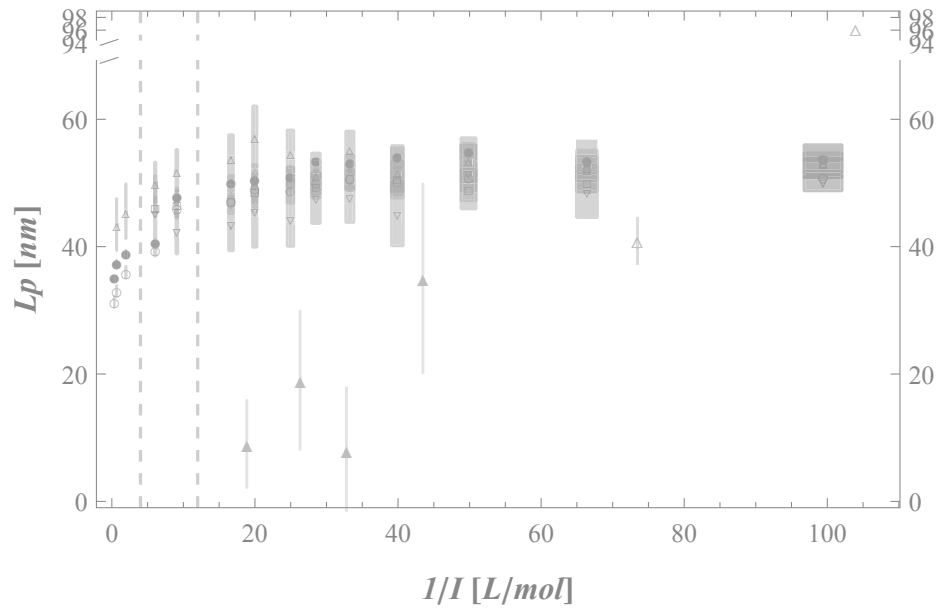


Fig. 5.11 – Comparaison de la variation de  $L_p$  obtenue pour  $L_{ADN} = 583, 639, 1201, 2060$  et  $3259$  bp sur notre gamme de force ionique induite par des ions  $\text{Mg}^{2+}$  avec les données de la littérature.

La comparaison de l'ensemble de ces résultats avec les nôtres ainsi que la discussion détaillée associée sont contenues dans l'article.

## 5.5 THÉORIE ODIJK, SKOLNICK, FIXMAN

Depuis le milieu des années 70, se construit une théorie à travers les travaux de **Odijk**, **Skolnick** et **Fixman**, qui fut nommée par la suite **théorie OSF** [Odi77, FS78]. Cette théorie propose une description des différents phénomènes physiques mis en jeu pour les **polyélectrolytes** en solution. Plus précisément, cette dernière exprime la dépendance de  $L_p$  en fonction des ions mobiles présents en solution, donc de la force ionique du milieu.

### 5.5.1 Longueur de persistance électrostatique

Le principe du calcul d'Odijk est de considérer l'énergie libre de la chaîne comme étant une somme de son énergie de courbure, et de l'énergie électrostatique induite par la présence des charges le long du polymère.

La longueur de persistance de la théorie OSF,  $L_p^{OSF}$  est alors définie comme la somme de la **longueur de persistance**  $L_p^\infty$ , correspondant à un ADN ayant toutes ses charges écrantées, soit  $A\kappa \rightarrow 0$  et de la **longueur de persistance électrostatique**  $L_{el}^{OSF}$ , due à la présence des charges sur l'ADN et des ions en solution. Ainsi, l'expression de  $L_p$  devient :

$$L_p^{OSF} = L_p^\infty + L_{el}^{OSF} \quad (5.15)$$

$$\Rightarrow L_p^{OSF} = L_p^\infty + \frac{l_B}{4A^2\kappa^2} \quad (5.16)$$

La longueur de persistance décrite est proportionnelle à l'inverse de la force ionique. L'expression de cette grandeur en fonction de la force ionique de la solution s'exprime donc :

$$L_p^{OSF} = L_p^\infty + \frac{l_B}{32A^2\pi l_B I_{[ion/m^3]}} \quad (5.17)$$

$$= L_p^\infty + \frac{1}{32\pi A^2 I_{[mol/L]} N_a \times 10^3} \quad [\text{nm}] \quad (5.18)$$

Pour un ADN double brin de forme B et de densité de charge linéique  $A_{ds}$ , l'expression de  $L_p^{OSF}$  devient :

$$L_p^{OSF} = L_p^\infty + \frac{0.559}{I} \quad [\text{nm}] \quad (5.19)$$

Ces relations sont valables lorsque la **longueur de persistance** est grande devant la **longueur de Bjerrum**, ce qui est le cas ici.

### 5.5.2 Limites de le théorie OSF

Cette théorie fut initialement développée en considérant le polymère tel un cylindre possédant une densité de charge linéique homogène en surface respectant

$L_p^{OSF} l_B \gg A^2$ . Le polymère est donc considéré comme un objet rigide. De plus, les charges portées le long du polyélectrolyte doivent rester relativement faibles afin que ces dernières puissent être décrites selon l'approximation de Debye-Hückel.

De plus, la théorie OSF est conçue afin d'exprimer la dépendance de la longueur de persistance à forte force ionique. La théorie initiale d'OSF ne prend pas en compte le phénomène de condensation énoncé plus haut.

### 5.5.3 Analyse de nos données par cette théorie

Les données de  $L_p$  sont ajustées par l'expression 5.19 avec deux paramètres libres, à forte force ionique induite par la présence d'ions  $Na^+$ , soit  $I \in [\sim 85 : 4000]$  mmol/L ou  $1/I < 12$  L/mol.

		Expression du fit : $L_p^{OSF} = L_p^\infty + \frac{C}{I}$					
$L_{ADN}$		575 bp	583 bp	639 bp	1201 bp	2060 bp	3259 bp
$Na^+$	$L_p^\infty$ , nm	52.8	43.3	43.6	31.7	35.5	45.2
	$C$ , nm.mol/L	1.44	1.00	5.75	1.52	1.21	1.61

L'ensemble des représentations associées à l'ajustement de nos données par la théorie OSF sont présentées en Fig. 5.12. Comme illustré, les expériences sur les ADN de longueurs  $L_{ADN} = 639$  et 3259 bp manquent de données pour que cet ajustement soit pertinent.

Quel que soit l'ADN étudié, l'estimation du paramètre  $C$  est éloignée du paramètre attendu d'après la théorie OSF. Nous obtenons des valeurs, en moyenne, 2 à 3 fois supérieures à la valeur théorique égale à  $C_{OSF} = 0.559$ .

L'estimation de  $L_p^\infty$  semble cohérente avec les connaissances de la physique du polymère ADN. La valeur moyenne obtenue est des  $40.0 \pm 6$  nm, (ou de  $42.0 \pm 7$  nm avec la prise en compte de  $L_{ADN} = 575$  bp).

Cette théorie ne permet pas une pleine analyse et compréhension des données.

## 5.6 PONDÉRATION À FAIBLE $I$ DE LA THÉORIE OSF PAR LE PARAMÈTRE DE MANNING

Afin d'étendre la compréhension de cette dépendance de  $L_p$  en fonction de  $I$ , la théorie de condensation des contre-ions présents en solution élaborée plus tôt par Manning est ajoutée à la théorie OSF, comme proposé dans les travaux de M. Le Bret et M. Fixman [Pod04, Fix82]. En effet, cette théorie est élaborée afin de décrire dans la limite des faibles force ioniques le phénomène de "condensation" des ions autour des charges présentes à la surface d'un polyélectrolyte.

La longueur de persistance s'exprime telle que :

$$L_p^{OSFM} = L_p^0 + L_{el}^{OSFM} \quad (5.20)$$

Le paramètre de pondération,  $\alpha_M$ , est inclus à l'expression de la **longueur de persistance électrostatique**, de la façon suivante :

$$\Rightarrow L_p^{OSFM} = L_p^0 + \frac{l_B \alpha_M^2}{4A_{ds}^2 \kappa^2} \quad (5.21)$$

Lorsque qu'il existe un phénomène de condensation des contre-ions, c'est-à-dire lorsque  $\xi > \frac{1}{|Z|}$ , alors  $\alpha_M$  est égal à l'inverse du paramètre de Manning  $\xi$  Eq. 5.13, ce dernier dépend de la nature des ions présent en solution. Dans notre cas, quelle que soit la nature des ions, la condensation des contre-ions présents en solution le long des charges du polymère s'applique.

Dans notre cas, les paramètres associés à la longueur de persistance OSF pondérée par le paramètre de condensation de Manning s'expriment tels que :

Nature des ions	<i>Monovalent</i> $Na^+$	<i>Divalent</i> $Mg^{2+}$
$\xi = Zl_B / A_{ds}$	4.2	8.4
Effet	$\xi > \frac{1}{ Z } \Rightarrow$ Condensation	
$\alpha_M = 1/\xi$	0.24	0.12
Expression	$L_p^{OSFM} = L_p^0 + \frac{l_B \alpha_M^2}{4A_{ds}^2 \kappa^2}$	
$L_p^{OSFM}$	$L_p^0 + \frac{0.033}{I}$	$L_p^0 + \frac{0.008}{I}$

### 5.6.1 Limites de la pondération

La remarque essentielle que nous pouvons déjà formuler vis-à-vis de cette approche est que cette dernière correspond à la combinaison d'une théorie élaborée pour le domaine de fort  $I$  avec la théorie de condensation de Manning, dans le but d'évaluer la dépendance de  $L_p$  à faible  $I$ . L'influence des forces ioniques intermédiaires, domaine de *transition*, sur le comportement du polymère n'est donc pas entièrement mise en équation.

### 5.6.2 Analyse de nos données par cette théorie

Les données de  $L_p$  sont ajustées par l'expression 5.21 avec un et deux paramètres libres, à forte force ionique induite par la présence d'ions  $Na^+$ , soit  $I \leq 85$  mmol/L ou  $1/I \in [12 : 100]$  L/mol.

Pour chaque nature d'ions, la première ligne de la table qui suit, correspond à l'ajustement avec un seul paramètre libre par l'expression Eq. 5.21. Le paramètre  $B$  est fixé à 0.033 dans le cas  $Na^+$  et 0.008 dans le cas  $Mg^{2+}$ . Les autres ajustements sont réalisés avec  $L_p^0$  et  $B$  comme paramètres libres. Les principaux ajustements sont proposés en Fig. 5.12.

		Expression du fit : $L_p^{OSFM} = L_p^0 + \frac{B}{I}$					
$L_{ADN}$		575 bp	583 bp	639 bp	1201 bp	2060 bp	3259 bp
$Na^+$	$L_p^0$ nm	77.5	52.4	45.9	48.5	50.9	48.7
$Na^+$	$L_p^0$ nm	69.0	51.6	44.8	48.5	51.0	48.7
	$B$ nm.mol/L	0.27	0.031	0.063	0.038	0.039	0.038
$Mg^{2+}$	$L_p^0$ nm	61.2	41.8	40.3	42.8	46.1	42.5
$Mg^{2+}$	$L_p^0$ nm	40.0	33.9	34.3	36.5	38.7	35.3
	$B$ nm.mol/L	0.36	0.14	0.14	0.10	0.18	0.14

### Gamme de $Na^+$

#### Observation sur l'estimation via un seul paramètre libre

La dépendance associée à cette estimation sur les données à faible force ionique obtenues en présence d'ions  $Na^+$  n'est pas satisfaisante pour rendre compte de la dépendance de  $L_p$  en fonction de  $I$ .

L'estimation de  $L_p^0$  semble cohérente et donne une valeur moyenne égale à  $49.3 \pm 3$  nm, (ou de  $54.0 \pm 12$  nm en tenant compte de l'ADN de 575 bp). Par contre, il est clair que le comportement à faible force ionique ne peut être extrapolé aux fortes forces ioniques.

#### Observation sur l'estimation avec deux paramètres libres

Afin de tester le poids de la valeur de  $B$ , nous avons ajusté nos données en laissant libre les deux paramètres. Les ajustements ne sont pas représentés.

L'estimation de la valeur de  $B$  est légèrement supérieure, mais en bon accord avec la valeur théorique attendue d'après la théorie OSFM. En effet, la valeur moyenne obtenue pour les cinq longueurs de molécule d'ADN, sauf l'ADN de 575 bp, qui nous avons étudiées est égale à  $0.042 \pm 0.008$  nm.mol/L, dont la barre d'erreur est égale à la déviation standard des données, comparable à la valeur théorique en présence d'ions monovalents égale à  $B = 0.033$  nm.mol/L, comme détaillé dans le tableau p169.

### Gamme de $Mg^{2+}$

#### Observation sur l'estimation avec un seul paramètre libre

La gamme de force ionique induite par la présence d'ions divalents  $Mg^{2+}$  conduit à une estimation de  $L_p^0$  de l'ordre de  $42.7 \pm 2$  nm, (ou de  $54.8 \pm 8$  nm avec la prise en compte de  $L_{ADN} = 575$ ). Cette valeur moyenne est plus faible que celle obtenue en présence des ions monovalents.

La qualité de l'ajustement est médiocre. Nous avons donc, pour la gamme en  $Mg^{2+}$ , ajusté la dépendance de  $L_p$  en fonction de  $I$ , en laissant deux paramètres libres, ces ajustements sont représentés en Fig. 5.12.

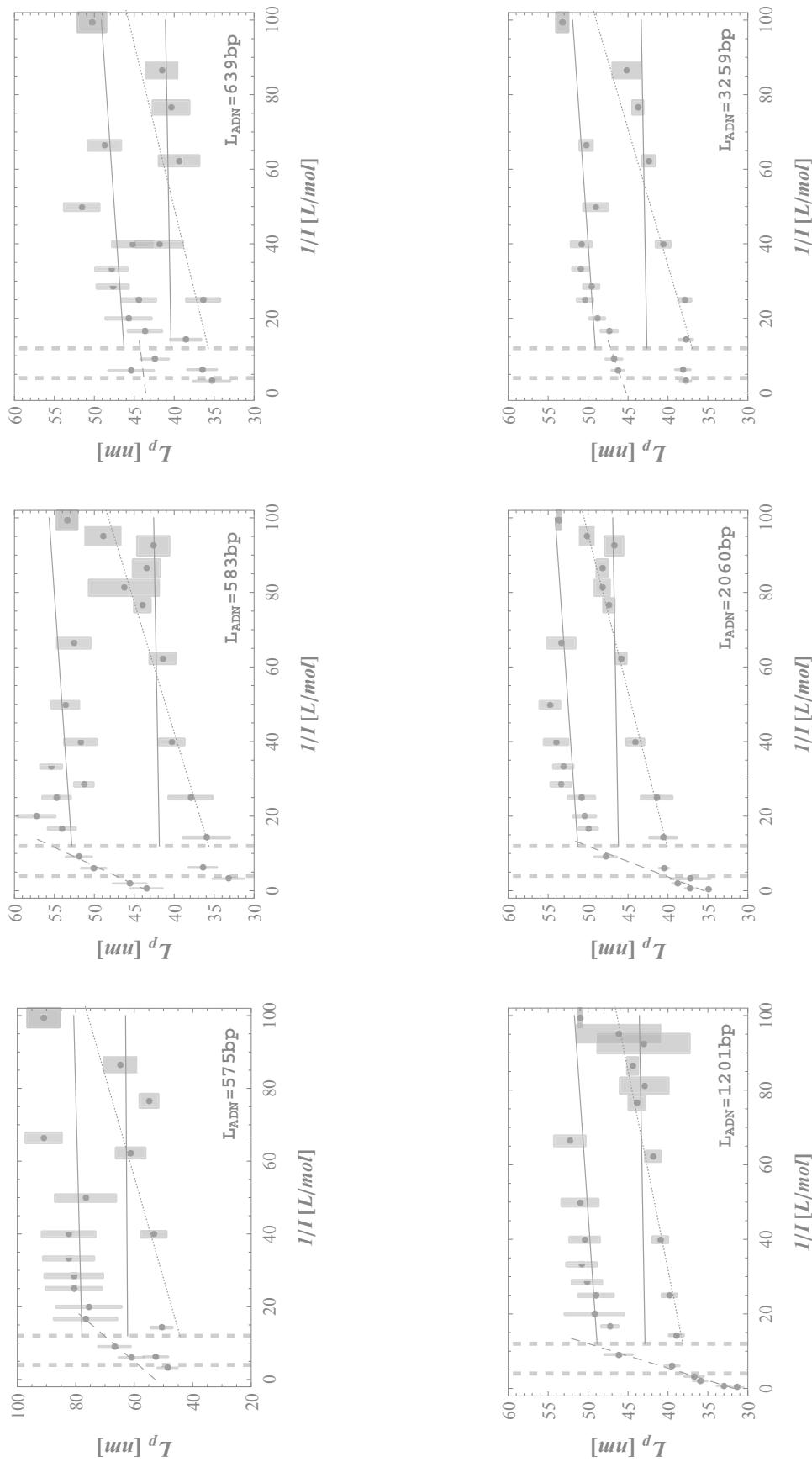


Fig. 5.12 – Variation de la longueur de persistance des ADN en fonction de la force ionique pour six longueurs d'ADN. Courbe en tirets, ajustement à forte force ionique en présence d'ions  $\text{Na}^+$  selon la théorie OSF, 5.19, valeur des paramètres en 5.5.3. Courbes en traits pleins, ajustement à faible force ionique en présence d'ions  $\text{Na}^+$  et  $\text{Mg}^{2+}$  selon la théorie OSFM, 5.21 avec un seul paramètre libre. Courbe en pointillés, sur la gamme d'ions  $\text{Mg}^{2+}$  selon la théorie OSFM, 5.21 avec deux paramètres libres. Valeur des paramètres en 5.5.3

### Observation sur l'estimation avec deux paramètres libres

Cette fois-ci, cet ajustement exprime assez bien la dépendance de  $L_p$  en fonction de  $I$ . L'estimation de  $L_p^0$  de l'ordre de  $35.7 \pm 2$  nm, (ou de  $36.4 \pm 2$  nm avec la prise en compte de  $L_{ADN} = 575$ ) pour une estimation de  $B$  de l'ordre de  $0.14 \pm 0.03$  nm.mol/L, (ou de  $0.18 \pm 0.09$  nm.mol/L avec la prise en compte de  $L_{ADN} = 575$ ).

L'estimation de la valeur de la longueur de persistance pour  $I \rightarrow 0$  est indéniablement plus basse en présence d'ions divalents que celle obtenue en présence d'ions monovalents. L'évaluation du paramètre  $B$  reste 2 fois supérieure à la valeur théorique attendue égale à  $B = 0.008$  nm.mol/L.

## 5.7 CORRECTION AUX VALEURS DE $I$ INTERMÉDIAIRES PAR UNE THÉORIE DE CHAMP MOYEN NON LINÉAIRE

Jusqu'à présent aucune théorie ne permet d'expliquer la concavité observée aux forces ioniques intermédiaires. Afin de prendre en compte cette variation de régime entre les faibles et les fortes forces ioniques, une approche consiste à appliquer une pondération de la densité de charge de l'ADN afin qu'elle soit dépendante de la force ionique du milieu. Cette idée fut notamment développée par Netz and Orland en 2003 [NO03].

Le paramètre  $\xi$  dépend alors de  $\kappa$ , autrement dit de  $I$ , et permet alors de pondérer le potentiel électrostatique existant en surface de l'ADN. Le détail du calcul est présenté dans la publication ainsi que dans son "*supplemental information*".

### 5.7.1 Dépendance de $\xi$ en fonction de $I$

Au vue de la gamme de forces ioniques prospectées (*correspondant aux symboles noirs sur la Fig. 5.13*), nous nous proposons d'utiliser une fonction d'interpolation polynômiale afin d'exprimer la dépendance de  $\xi_{Na^+}$  en fonction des grandeurs d'écrantage (*correspondant à la courbe épaisse rouge sur la Fig. 5.13*).

En présence d'ions  $Mg^{2+}$ , la loi de puissance suivante (*correspondant à la droite fine bleue sur la Fig. 5.13*), permet de rendre compte très convenablement de la dépendance de  $\xi_{Mg^{2+}}$  (*correspondant à la courbe épaisse bleue sur la Fig. 5.13*), uniquement sur la gamme en sel explorée. En effet, si nous avons sondé une gamme de force ionique plus importante, notamment sur les fortes forces ionique, l'erreur induite par cette approximation est alors de l'ordre de 10%.

$$\xi_{Mg^{2+}}(\kappa R_{ADN}) = 0.424(\kappa R_{ADN})^{0.364} \quad (5.22)$$

L'approximation de la dépendance de  $\xi_{Na^+}$  par une telle loi ne permettrait pas de rendre compte du comportement de cette dernière à forte force ionique (*les trois derniers marqueurs noirs sur la Fig. 5.13*).



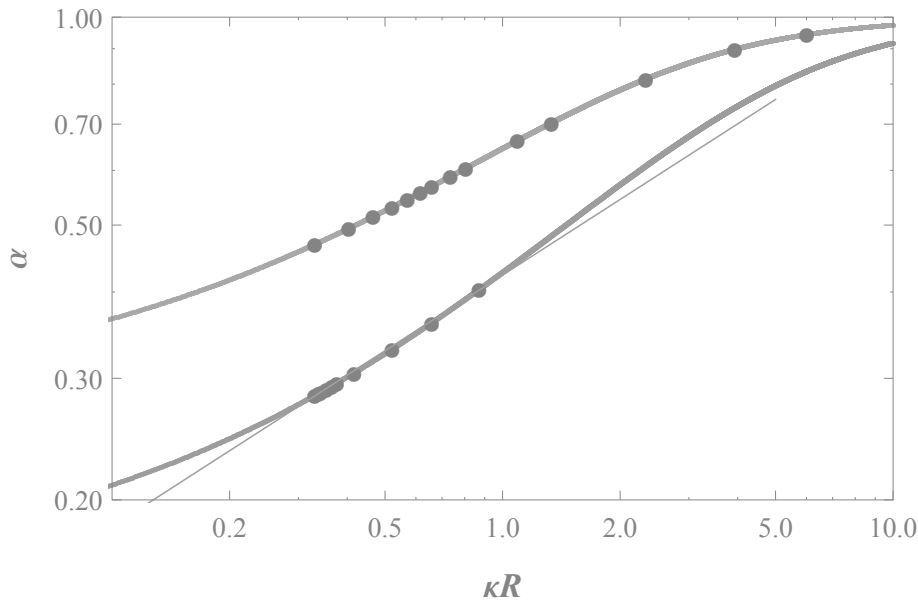


Fig. 5.13 – Représentation de la fonction de pondération de la densité de charge en fonction des grandeurs d'écrantage

L'expression de la longueur de persistance est alors définie par :

$$L_p^{MF} = L_p^0 + \frac{l_B}{4 A^2 \kappa^2} \cdot \alpha_{ion}^2(\kappa \cdot R_{ADN}) \quad (5.23)$$

### 5.7.2 Limites de la théorie de champ moyen

Ici, l'équation de Poisson-Boltzmann est approchée par l'équation de Debye-Hückel, avec le paramètre variationnel  $\alpha$ , afin d'exprimer le potentiel électrostatique et la densité de charge de l'ADN. Cette approche de type champ moyen ne prend en compte aucune corrélation entre les charges. De plus, là encore, la structure de l'ADN est considérée comme complètement impénétrable par les ions. Autrement dit les charges sont considérées comme restant en surface de l'ADN modélisé par un cylindre chargé, ce qui n'est pas une vision entièrement réaliste.

### 5.7.3 Analyse de nos données par cette théorie

Les données de  $L_p$  sont ajustées par l'expression 5.23 avec deux paramètres libres,  $L_p^0$  et  $D$  un pré-facteur du deuxième terme de cette équation, sur l'ensemble de la gamme de  $I$ .

		Expression du fit : $L_p^{MF} = L_p^0 + D \alpha_{ion}^2 \frac{l_B}{4 \times A^2 \kappa^2}$					
$L_{ADN}$		575 bp	583 bp	639 bp	1201 bp	2060 bp	3259 bp
$Na^+$	$L_p^0$ , nm	64.7	49.1	42.7	44.2	45.9	46.3
	$D$	2.39	0.57	0.69	0.80	0.91	0.57
$Mg^{2+}$	$L_p^0$ , nm	43.0	31.7	34.0	35.2	35.7	34.3
	$D$	2.5	1.86	1.54	1.54	1.79	1.72

Les représentations associées à ces ajustements sont présentées en Fig. 5.15.

Le facteur numérique  $D$  associé à la pondération de la densité de charge portée par l'ADN n'est pas suffisant pour expliquer nos données expérimentales. En effet, en présence d'ions monovalents sa valeur moyenne est de  $0.71 \pm 0.15$ . Il semblerait donc nécessaire de diminuer l'expression théorique d'environ un facteur  $\frac{3}{4}$ . La valeur moyenne obtenue pour  $L_p^0$  est de l'ordre de  $45.6 \pm 2$  nm (ou de  $48.8 \pm 8$  nm avec la prise en compte de  $L_{ADN} = 575$ ).

En présence d'ions divalents, il est nécessaire d'augmenter la valeur de ce préfacteur, dont la valeur moyenne est de l'ordre de  $1.69 \pm 0.15$ . Ainsi, la décroissance de  $L_p$  est plutôt bien expliquée par cette dernière loi. La valeur moyenne obtenue pour  $L_p^0$  reste toujours inférieure en présence d'ions divalents, elle est de l'ordre de  $34.2 \pm 2$  nm, (ou de  $35.7 \pm 5$  nm avec la prise en compte de  $L_{ADN} = 575$ ).

Par contre en présence d'ions monovalents, là encore, cette théorie ne suffit pas à exprimer la brusque chute observée à forte force ionique. Ceci peut s'expliquer par l'absence de corrélations dans le modèle, ce qui est discuté dans la publication.

## 5.8 PRISE EN COMPTE DES FORCES D'ÉTIREMENT DANS UNE NOUVELLE THÉORIE DE MANNING

Fort de cette nécessité d'exprimer la dépendance de la longueur de persistance sur l'ensemble de la gamme de force ionique qui peut être explorée expérimentalement, Manning élabore une nouvelle théorie en 2006 [Mano6].

Cette nouvelle approche prend en compte l'effet de la répulsion électrostatique existante entre les charges le long du polymère sur l'étirement local de la chaîne. Or il a été mis en lumière que les charges négatives portées par le squelette de l'ADN mettent en jeu de fortes énergies de répulsion. Cette répulsion conduit à une force d'étirement intrinsèque de la molécule d'ADN, qui n'est pas négligeable à faible force ionique [Neto1]. Cette approche permet donc de prendre en compte les interactions qui tendent à aligner les monomères du polymère mais aussi les forces qui tendent à les étirer, Fig. 5.14.



Fig. 5.14 – Représentation schématique des forces d'interaction prise en compte dans les modèles précédents, à gauche, dans la nouvelle théorie de Manning, à droite. Les flèches bleues représentent les interactions qui tendent à aligner le polymère, les rouges celles qui tendent à l'allonger.

Cette théorie est basée sur le développement présenté dans la partie précédente auquel est associé la longueur de persistance de l'*isomère virtuel sans charge* notée  $L_p^*$ . Cet *isomère virtuel sans charge* correspond à l'hypothèse où aucun groupement phosphate de l'ADN n'est ionisé,  $L_p^*$  est donc associée à la grandeur caractérisant un polyélectrolyte dont l'ensemble de charges portées le long de son squelette seraient mises à zéro.

Ainsi la longueur de persistance s'exprime selon l'expression suivante [Mano6] :

$$L_p^{MF} = \left(\frac{\pi}{2}\right)^{2/3} \frac{R_{ADN}^{4/3} (L_p^*)^{2/3}}{Z^2 \ell_B} \left[ (2\xi - 1) \frac{\kappa A e^{-\kappa A}}{1 - e^{-\kappa A}} - 1 - \ln(1 - e^{-\kappa A}) \right] \quad (5.24)$$

### 5.8.1 Analyse de nos données par cette théorie

Les données de  $L_p$  sont ajustées par l'expression 5.24 avec un seul paramètre libres,  $L_p^*$ , sur l'ensemble de la gamme de  $I$ .

		Expression du fit : $L_p^{MF} = \left(\frac{\pi}{2}\right)^{2/3} \frac{R_{ADN}^{4/3} (L_p^*)^{2/3}}{Z^2 \ell_B} \left[ (2\xi - 1) \frac{\kappa A e^{-\kappa A}}{1 - e^{-\kappa A}} - 1 - \ln(1 - e^{-\kappa A}) \right]$					
$L_{ADN}$		575 bp	583 bp	639 bp	1201 bp	2060 bp	3259 bp
$Na^+$	$L_p^*$ nm	11.3	6.4	5.1	5.7	6.0	5.7
$Mg^{2+}$	$L_p^*$ nm	22.0	12.6	12.0	12.7	14.2	13.8

Les représentations associées à ces ajustements sont présentées en Fig. 5.15. Cette approche produit les meilleurs ajustements de nos données réalisés jusqu'à présent, sur l'ensemble de la gamme de force ionique réalisée en présence d'ions monovalents. En effet, cette théorie permet d'exprimer le profil de  $L_p$  observé à faible  $I$  tout en évaluant correctement la brusque chute de la longueur de persistance à forte force ionique. La valeur moyenne obtenue pour  $L_p^*$  en présence d'ions monovalents est de l'ordre de  $5.8 \pm 0.5$  nm (ou de  $6.7 \pm 2$  nm avec la prise en compte de  $L_{ADN} = 575$ ).

L'ajustement associée à la gamme en sels divalents n'est cependant pas satisfaisante. En effet, la diminution de  $L_p$  dans la gamme de force ionique intermédiaire est plus importante dans nos données que celle évaluée via le fit. De plus, la valeur moyenne obtenue pour  $L_p^*$  est près de 2 fois supérieure en présence d'ions divalents, de l'ordre de  $13 \pm 1$  nm (ou de  $14 \pm 4$  nm avec la prise en compte de  $L_{ADN} = 575$ ). Tout ceci souligne une fois encore, la différence de comportement de la réponse de l'ADN en présence d'ions monovalents et en présence d'ions divalents. La prise en compte de cette différence via le paramètre de Manning n'est pas suffisante. À ce jour, de nouvelles théories sont donc encore nécessaires.

L'ajustement de  $L_p^*$  obtenu sur l'ensemble des tailles de molécules d'ADN, ne nous permet pas d'exprimer une corrélation entre les paramètres  $L_p^*$  et  $L_{ADN}$ , que se soit lors de l'évaluation de sa dépendance en présence d'ions monovalent  $Na^+$  ou en présence d'ions divalents  $Mg^{2+}$ .

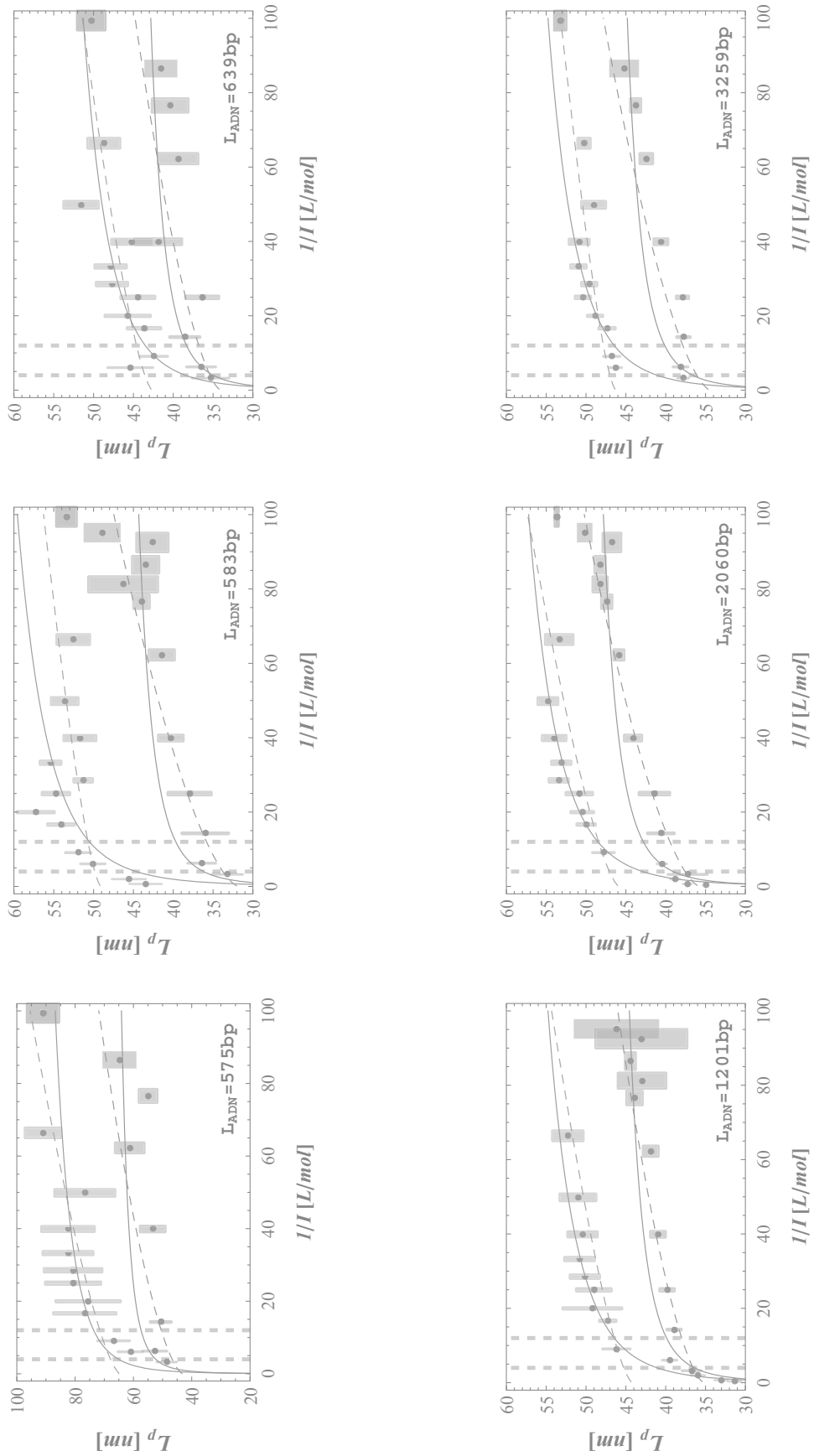


Fig. 5.15 – Variation de la longueur de persistance des ADN en fonction de la force ionique pour six longueurs d'ADN. Courbe en tirets, ajustement, sur toute la gamme de concentration en ions  $\text{Na}^+$  et  $\text{Mg}^{2+}$ , selon la théorie 5.23 avec deux paramètres libres, valeur des paramètres en 5.7.3. Courbes en traits pleins, ajustement, sur tout la gamme de force ionique, selon la théorie nouvelle théorie de Manning, 5.24, avec un seul paramètre libre. Courbe en pointillés, sur la gamme d'ions  $\text{Mg}^{2+}$  avec un seul paramètre libre, valeur des paramètres en 5.8.1

## 5.9 PROPOSITION D'UNE FORMULE D'INTERPOLATION

Afin de rendre compte de la dépendance de  $L_p$  sur l'ensemble de notre gamme de concentration en sel, ainsi que sur l'ensemble de la nature des ions testés nous proposons la formule d'interpolation phénoménologique suivante :

$$L_p = L_p^\infty + \frac{L_p^0 - L_p^\infty}{1 + (I/I_0)^\delta} \quad (5.25)$$

Cette dernière rend compte de la différence de longueurs de persistance obtenue à fort sel et à faible sel. En effet, lorsque  $I \rightarrow \infty$  alors  $L_p \rightarrow L_p^\infty$ , suivant la théorie OSF. Tandis que lorsque  $I \rightarrow 0$  alors  $L_p \rightarrow L_p^0$  suivant l'apport de la combinaison des théories d'OSF et de Manning.

### 5.9.1 Analyse de nos données par cette théorie

Les données de  $L_p$  sont ajustés par l'expression 5.25, avec quatre paramètres libres,  $L_p^\infty$ ,  $L_p^0$ ,  $I_0$  et  $\delta$ , sur l'ensemble de la gamme de concentration en sel monovalent.

Ensuite, la valeur de  $L_p^\infty$  est fixée à  $L_{p,Na^+}^\infty$ . Ceci est justifié dans le cas où nous prospectons de plus fortes forces ioniques avec les ions  $Na^+$  qu'avec les ions  $Mg^{2+}$ . Néanmoins, nous réalisons un second ajustement sur la gamme  $Mg^{2+}$  avec l'ensemble des 4 paramètres libres.

		Expression du fit : $L_p = L_p^\infty + \frac{L_p^0 - L_p^\infty}{1 + (I/I_0)^\delta}$					
$L_{ADN}$		575 bp	583 bp	639 bp	1201 bp	2060 bp	3259 bp
$Na^+$	$L_p^\infty$ , nm	102.7	54.5	58.5	53.9	58.3	57.6
	$L_p^0$ , nm	13.2	38.5	26.9	30.1	33.2	27.4
	$I_0$ , mol/L	0.218	0.560	0.188	0.174	0.137	0.607
	$\delta$	0.600	0.998	0.318	0.994	0.735	0.360
$Mg^{2+}$	$L_p^0$ , nm	51.7	63.6	51.1	51.7	59.0	58.5
	$I_0$ , mol/L	0.045	0.004	0.048	0.045	0.018	0.0217
	$\delta$	0.546	1.000	0.501	0.546	0.755	0.393
$Mg^{2+}$	$L_p^\infty$ , nm	77.7	52.9	52.2	50.8	64.9	53.0
	$L_p^0$ , nm	38.6	27.2	23.0	30.6	31.2	33.8
	$I_0$ , mol/L	0.038	0.035	0.082	0.040	0.013	0.022
	$\delta$	0.906	0.724	0.412	0.624	0.593	0.718

Cette formule phénoménologique nous permet de décrire correctement la dépendance observée de  $L_p$  en fonction de la force ionique induite par la présence d'ions monovalents ainsi que divalents, Fig. 5.16.

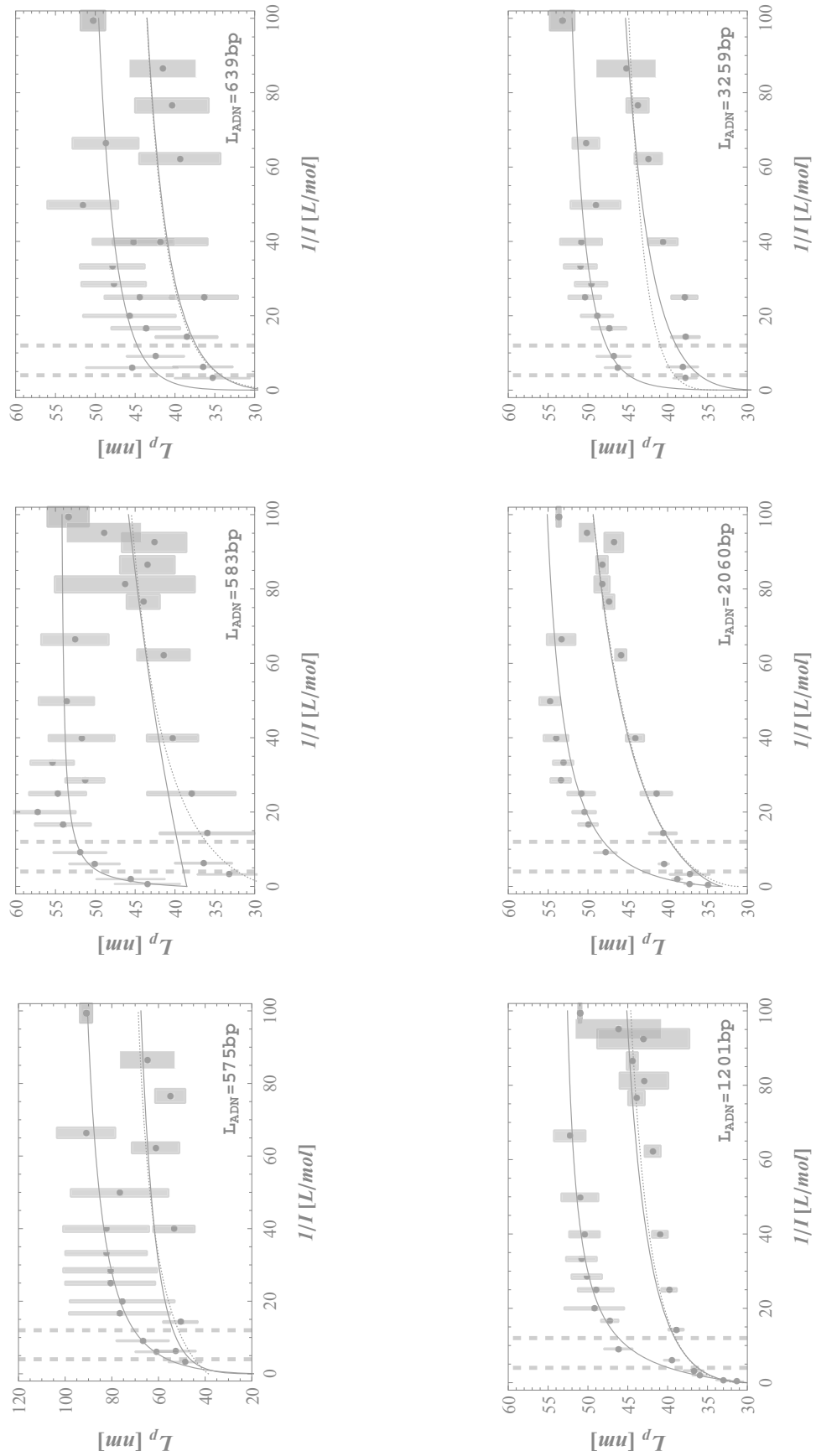


Fig. 5.16 – Variation de la longueurs de persistance des ADN en fonction de la force ionique pour six longueurs d'ADN. Courbes en trait plein, ajustement, sur toute le gamme de concentration en ions  $\text{Na}^+$ , selon la formule d'interpolation 5.25 avec quatre paramètres libres. Courbes en trait pointillé, ajustement, sur toute le gamme de concentration en ions  $\text{Mg}^{2+}$ , selon cette même formules, avec  $L_p^\infty = L_{p,\text{Na}^+}^\infty$  et trois paramètres libres (ce qui explique la faible qualité de l'ajustement pour l'ADN de 583 bp). Les courbes en pointillés correspondent à une évaluation sur toute la gamme de concentrations en ions  $\text{Mg}^{2+}$  avec quatre paramètres libres. L'ensemble des valeurs des paramètres est listé en 5.8.1

Sans considérer l'ADN de 575 bp, intéressons-nous aux moyennes des valeurs extraites pour les différents paramètres libres. Pour cela nous nous focalisons sur les meilleures évaluations de la gamme  $Mg^{2+}$  obtenues avec un ajustement avec 3 paramètres libres, sauf pour l'ADN  $L_{ADN} = 583$  bp, pour lequel nous regarderons celui obtenu avec 4 paramètres libres.

Suite à l'analyse, la valeur moyenne de  $L_{p,Na^+}^{\infty} = 55.6 \pm 2$  nm. Cette valeur est comparable à l'estimation de  $L_{p,Mg^{2+}}^0 = 54.6 \pm 3$  nm.

La valeur moyenne de l'estimation de  $I_0$  est égale à  $I_{0,Na^+} = 0.333 \pm 0.2$  sur la gamme en  $Na^+$ . Cette valeur est dix fois inférieure sur la gamme en  $Mg^{2+}$   $I_{0,Mg^{2+}} = 0.033 \pm 0.02$  mol/L.

L'ajustement du paramètre  $\delta$ , n'aboutit pas à une bonne précision de sa valeur. Cependant la valeur moyenne obtenue sur la gamme en  $Na^+$  est de  $\delta = 0.62 \pm 0.3$  et de  $\delta = 0.68 \pm 0.2$  sur la gamme en  $Mg^{2+}$ , ce qui conduit à une valeur moyenne de  $\delta = 0.65 \pm 0.4$  pour les deux natures d'ions confondues.

Cette formule phénoménologique peut permettre l'interpolation de la valeur de  $L_p$  sur l'ensemble de la gamme de force ionique prospectée et ce, quelle que soit la nature des ions présents, mais elle n'a pas de justification physique à ce jour.

## 5.10 MESURE D'UN EFFET SÉQUENCE PAR L'APPROCHE TPM

Avant de conclure ce chapitre revenons sur la différence de comportement que nous avons pu observer pour l'ADN 575. Dès la mesure expérimentale de la distance bout-à-bout des complexes ADN/particule, ce dernier se singularise. En effet, les valeurs de  $R_{Exp||}$  sont nettement plus élevées pour  $L_{ADN} = 575$  bp que pour  $L_{ADN} = 583$  bp, se rapprochant plus des mesures obtenues pour  $L_{ADN} = 639$  bp. Quelle que soit alors la méthode d'extraction, l'estimation de  $L_p$  répercute cette particularité.

Ceci nous amène à penser que l'ADN de 575 bp possède une spécificité de séquence qui influe sur ses propriétés physiques. Afin de prospecter l'impact de la séquence ADN sur l'influence de la force ionique sur  $L_p$ , nous nous sommes intéressés à deux ADN de même taille, mais différant seulement sur des séquences dont les positions sont limitées et connues.

### 5.10.1 Variation de la distance bout-à-bout des complexes ADN/particule

Nous considérons deux constructions de même longueur :

- $L_{ADN} = 583$  bp dont la séquence peut être considérée comme aléatoire, correspondant à l'ADN dont les résultats furent jusqu'à présent présentés.
- $L_{ADN,Poly(dA)} = 583$  bp, copie de la molécule précédente, sauf sur 60 bp au milieu de la molécule qui sont remplacés par un  $Poly(dA)$ . Un  $Poly(dA)$  est une séquence d'ADN uniquement constituée de bases **A** interagissant spécifiquement avec les bases **T** du brin complémentaire.

Ces spécificités structurales peuvent impacter la réponse à différentes concentrations en sel. En effet, comme nous l'avons vu précédemment le couple G-C est associé à une plus grande zone de charge, par opposition au couple A-T, Fig. 1.14. De plus, dès 1987 des études ont montré que la double hélice d'ADN de forme B incluant des séquences *Poly(dA)* ou *Poly(dT)* présentent des propriétés structurales différentes [PW81, ALS, SGP05]. En effet, les séquences *Poly(dA)* ou *Poly(dT)* présentent un petit sillon plus étroit.

Explorer la réponse induite par la présence d'un *Poly(dA)* à différentes concentrations en sel, peut nous permettre de mettre en évidence des différences de comportement induites par la nature et la séquence même des bases azotées.

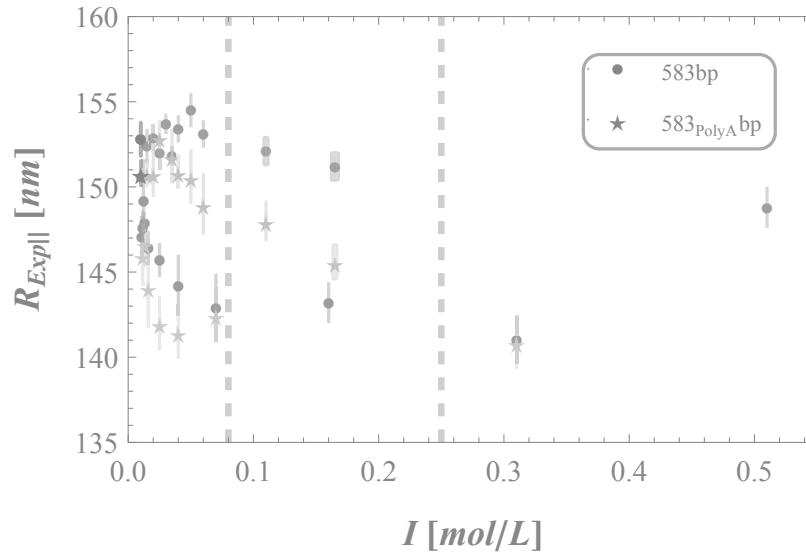


Fig. 5.17 – Comparaison de la distance bout-à-bout obtenue pour un ADN de 583 bp contenant ou non un *Poly(dA)* de 60 bp

Sur la Fig. 5.17 la valeur de  $R_{Exp||}$  pour  $L_{ADN, Poly(dA)} = 583$  bp est systématiquement inférieure à celle pour  $L_{ADN} = 583$ . Intéressons-nous à la variation de la longueur de persistance, Fig. 5.18.

### 5.10.2 Variation de la longueur de persistance

Les valeurs de  $L_p$  sont extraites via la résolution du problème inverse, basée sur les simulations par échantillonnage exact d'une molécule d'ADN de 583 bp avec  $N = 49 + 1$  et  $R_{par} = 150$  nm, Partie 5.2.3. Ces dernières sont représentées sur la Fig. 5.18. L'effet est significatif sur les concentrations intermédiaires en ions monovalents  $Na^+$ . Sur cette plage, la chute des  $L_p$  apparaît pour des valeurs de force ionique plus faibles pour  $L_{ADN, Poly(dA)} = 583$  bp que pour  $L_{ADN} = 583$  bp.

Les conformations de l'ADN présentant une séquence *Poly(dA)* semblent plus sensibles à forte force ionique. Les bases G-C présentent de forts domaines de charge, négatifs et positifs, dans le grand sillon, Partie 1.14. Le potentiel électrostatique associé peut contribuer à limiter la condensation des contre-ions le long de l'ADN. À l'inverse, une séquence *Poly(dA)* encombrera moins le paysage électrostatique associé à la molécule, tel un défaut de charge, ceci peut induire une zone favorable à une plus grande amplitude du phénomène de condensation et donc à une réduction de  $L_p$ .



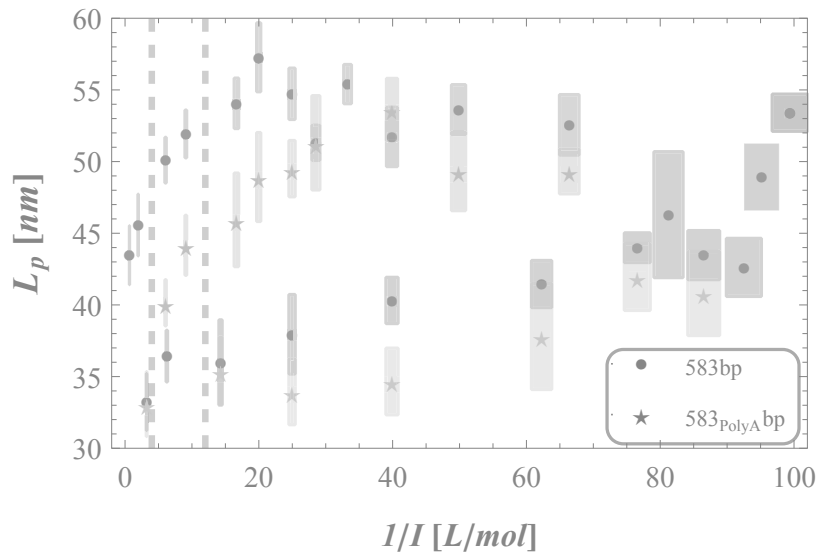


Fig. 5.18 – Comparaison de la longueur de persistance obtenue pour un ADN de 583 bp contenant ou non un Poly(dA) de 60 bp.

### 5.10.3 Ajustement via la nouvelle théorie de Manning

Nous utiliserons seulement la nouvelle théorie de Manning afin d'évaluer les variations des propriétés physiques intrinsèques aux 2 molécules d'ADN, Fig. 5.19. Pour une même longueur d'ADN, l'écart entre ces deux séries de mesures est notable. L'ajustement via la nouvelle théorie de Manning, via 5.24 de la Partie. 5.8, souligne une différence de  $\sim 15\%$  dans l'estimation des longueurs de persistance nue,  $L_p^*$ , et ce quelle que soit la nature des ions. Cet écart semble donc correspondre à une modification des propriétés physiques intrinsèques à la molécule d'ADN par la présence d'une séquence Poly(dA).

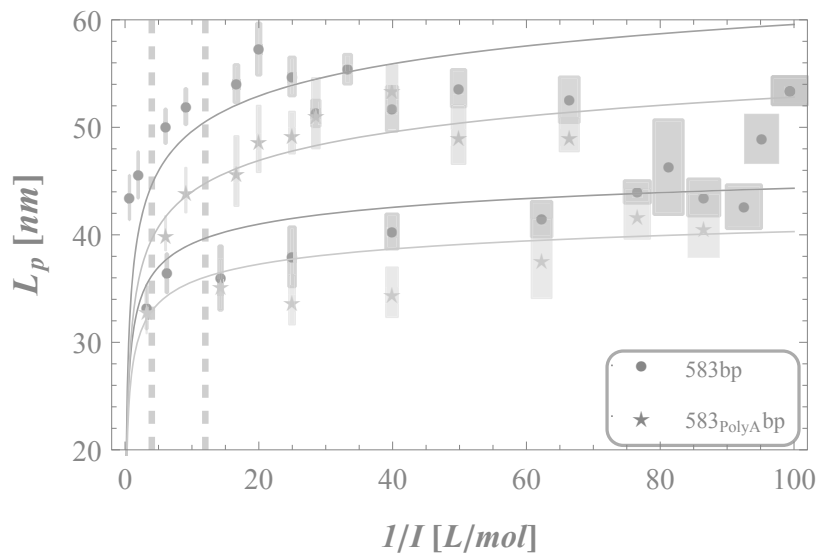


Fig. 5.19 – Comparaison de la longueur de persistance obtenue pour un ADN de 583 bp contenant ou non un Poly(dA) de 60 bp. Courbes en trait plein rouge et orange, ajustement, sur toute la gamme de concentration en ions  $\text{Na}^+$ , selon la formule d'interpolation 5.25 avec quatre paramètres libres. Courbes en trait plein bleu et bleu clair, ajustement, sur toute le gamme de concentration en ions  $\text{Mg}^{2+}$ , selon cette même formule, avec trois paramètres libres.

Par extrapolation, à la limite inférieure de la gamme de conditions salines physiologiques, la variation relative de  $L_p$  entre ces deux ADNs est de 16% en présence d'ions monovalent, de 42 à 50 nm, et de 23 % en présence d'ions divalents, de 33 à 40 nm.

		Expression du fit : $L_p^{MF} = \left(\frac{\pi}{2}\right)^{2/3} \frac{R_{ADN}^{4/3} (L_p^*)^{2/3}}{Z^2 \ell_B} \left[ (2\zeta - 1) \frac{\kappa A e^{-\kappa A}}{1 - e^{-\kappa A}} - 1 - \ln(1 - e^{-\kappa A}) \right]$	
		$L_{ADN} = 583 \text{ bp}$	$L_{ADN, Poly(dA)} = 583 \text{ bp}$
$Na^+$	$L_p^*$	6.4	5.4
$Mg^{2+}$	$L_p^*$	12.6	10.4

Des premières mesures ont permis de montrer que la longueur de persistance associée à un *Poly(dA)*, ou *Poly(dT)*, est d'environ 30 – 40% inférieure à celle d'un *Poly(dG)* ou *Poly(dC)* [HLA83]. Peu après, des mesures de cyclisation ont été réalisées en présence de séquences courtes de 6 bp de *Poly(dA)* répétées représentant 50% de la séquence totale. Cette dernière étude retrouve une diminution de  $L_p$ , moins importante que précédemment, sur les molécules présentant des *Poly(dA)* [UBTC86].

Les séquences *Poly(dA)* ou *Poly(dT)* présentent un véritable intérêt biologique. En effet, un certain nombre de régions promoteur de gènes<sup>1</sup> sont délimitées par ce type de séquences [SR90, VCDC96]. Les caractéristiques de telles séquences, ici en fonction des gradients de concentration qui peuvent intervenir lors de certains processus biologiques, est un atout dans la compréhension des propriétés physiques et mécaniques de ce type de séquences ainsi que dans leur rôle possible *in vivo*.

Cette étude montre qu'une séquence *Poly(dA)* influe sur la mesure de la longueur apparente de la molécule d'ADN. Cette variation se répercute sur l'estimation de la longueur de persistance. Seules, ces mesures ne peuvent nous permettre de définir la nature même de l'effet observé. Une variation locale dans la molécule d'ADN telle que la présence d'un défaut de charge, d'une courbure locale ou encore l'ouverture d'une bulle de dénaturation ne peut, en l'état actuel, être facilement différenciée d'une modification locale de rigidité. Pour cela, des études supplémentaires sont nécessaires afin de nous permettre de discriminer les réponses associées à chacun de ces effets. Par contre, la présence d'une séquence *Poly(dA)* influe sur la réponse physique et mécanique de la molécule en fonction de la concentration et la nature des ions présents en solutions.

Si nous revenons à l'ADN de 575 bp, il semble donc probable que nos mesures soulignent un effet séquence. Celui-ci reste pour l'instant mal compris. En effet, cette molécules d'ADN ne présente pas des singularités ou des séquences spécifiques particulières. De plus amples expériences semblent donc nécessaires afin de comprendre la nature du phénomène qui induit des propriétés physiques intrinsèques particulières à ce polymère. Plus précisément d'où provient cette augmentation de la rigidité de la molécule d'ADN de 575 bp ?

1. Tel que les gènes : HIS3, PET56, DED1, CBSA, ARG4 ou encore UR43

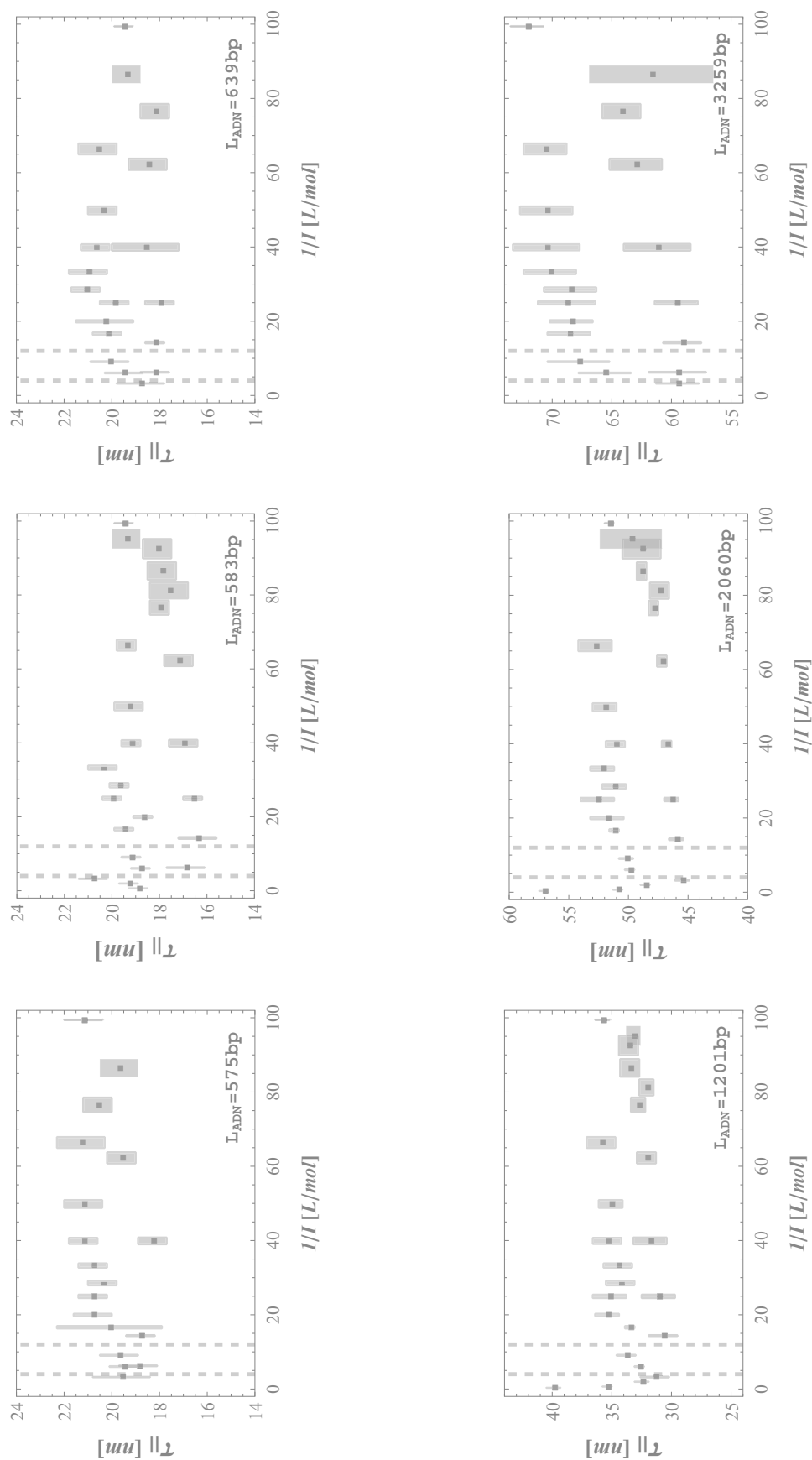


Fig. 5.20 – Variation du temps de relaxation des complexes ADN/particules en fonction de la force ionique pour six longueurs d'ADN. Le même code couleur que précédemment est respecté, soit rouge pour les mesures en présence d'ion monovalents  $\text{Na}^+$ , bleu pour les mesures en présence d'ions divalents  $\text{Mg}^{2+}$ , et noir les mesures à Zero-Salt-Buffer

### 5.11 OBSERVATION SUR LA VARIATION DU TEMPS DE RELAXATION

Pour finir commentons le fait que le temps de relaxation associé à nos complexes ADN/particule varie en fonction de la force ionique, Fig. 5.20. En effet, ce dernier semble diminuer lentement et progressivement en fonction de l'augmentation de la concentration en sel. Quelques travaux dans la littérature rendent compte de cette même observation, se référer à l'annexe de l'article pour plus de détails.

Une explication de ce phénomène est qu'une diminution de  $L_p$  induit une diminution du rayon hydrodynamique de la molécule d'ADN et donc ceci implique une décroissance du temps de relaxation,  $\tau_{||}$ , associé au système.

Soulignons toutefois que la valeur absolue de  $\tau_{||}$  en présence d'ions divalents  $Mg^{2+}$  reste incontestablement inférieure à sa mesure en présence d'ions monovalents. De plus, à très forte force ionique, distinctement observable seulement pour  $L_{ADN} = 1201$  et  $2060$  bp en présence d'ions  $Na^+$ ,  $\tau_{||}$  augmente abruptement.

## CONCLUSION DU CHAPITRE

Comprendre les effets de l'environnement sur les propriétés mécaniques de l'ADN est essentiel pour mieux appréhender ce qui se produit à l'intérieur du noyau cellulaire, notamment lors de processus biologiques tels que la transcription ou la réplication, ou encore dans les capsides virales. Les molécules chargées présentes en solution, que ce soit des ions ou des protéines, sont les partenaires privilégiés de la molécule d'ADN. La présence d'ions de différentes natures à différentes concentrations, influe donc directement sur les propriétés physiques du biopolymère.

De plus, de nombreuses études soulignent un effet des ions et contre-ions sur la caractéristique intrinsèque de l'ADN qu'est la longueur de persistance. La modification de cette grandeur se répercute sur la longueur apparente de la molécule. La technique TPM de molécule unique est donc idéalement adaptée pour sonder ce type de fluctuation, sans contraindre les degrés de libertés de la molécule par l'application de forces extérieures.

Nous avons donc réalisé des mesures sur une gamme riche aux forces ioniques faibles et intermédiaires, et étendue jusqu'aux fortes forces ioniques. Tout d'abord, ceci nous a permis de relever, comme attendu, une forte dépendance de la distance bout-à-bout de nos complexes ADN/particule, puis d'extraire la variation de la longueur de persistance en fonction de  $I$ .

Lors de l'extraction de  $L_p$  de nos mesures expérimentales, nous avons pu juger de l'erreur induite par les méthodes de soustraction de la contribution de la bille ainsi que des méthodes d'extraction de  $L_p$  basées sur les modèles physiques des polymères. Utilisées de façon brute, ces dernières ne prennent pas bien en compte les caractéristiques et contraintes liées aux expériences. La réalisation et l'utilisation d'un modèle numérique simulant ces dernières, permet de pallier certaines erreurs et approximations.

Une fois  $L_p$  ainsi extraite, nous avons pu pleinement évaluer et analyser la variation de cette grandeur en fonction de la concentration des contre-ions en solution. Cette dépendance diffère en fonction de la nature des ions, monovalents  $Na^+$  ou divalents  $Mg^{2+}$ , présents en solution.

Dans la littérature la variation de  $L_p$  en fonction de  $I$  induite par la présence d'ions monovalents est encore sujette à controverse. En effet, les résultats expérimentaux peuvent être divisés en deux grands groupes définis selon le profil de  $L_p$  à forte force ionique. Le premier groupe observe une lente variation de  $L_p$  lorsque  $I$  augmente dans cette gamme de forte concentration en ions. Tandis que le second groupe observe une variation très significative. Nos données sont en accord avec ce dernier groupe pour la gamme de forte force ionique. Par contre, nous obtenons une variation de  $L_p$  aux faibles forces ioniques qui leur est inférieure, mais en accord avec la variation observée par le premier groupe. Le manque de données en présence d'ions divalents  $Mg^{2+}$  ne nous permet pas une comparaison rigoureuse. Tout ceci fait écho au désaccord actuel sur les gammes d'influence et l'amplitude de la dépendance de  $L_p$  en fonction de  $I$ . De plus, ceci souligne la difficulté à laquelle se heurtent les scientifiques afin d'élaborer les théories associées.

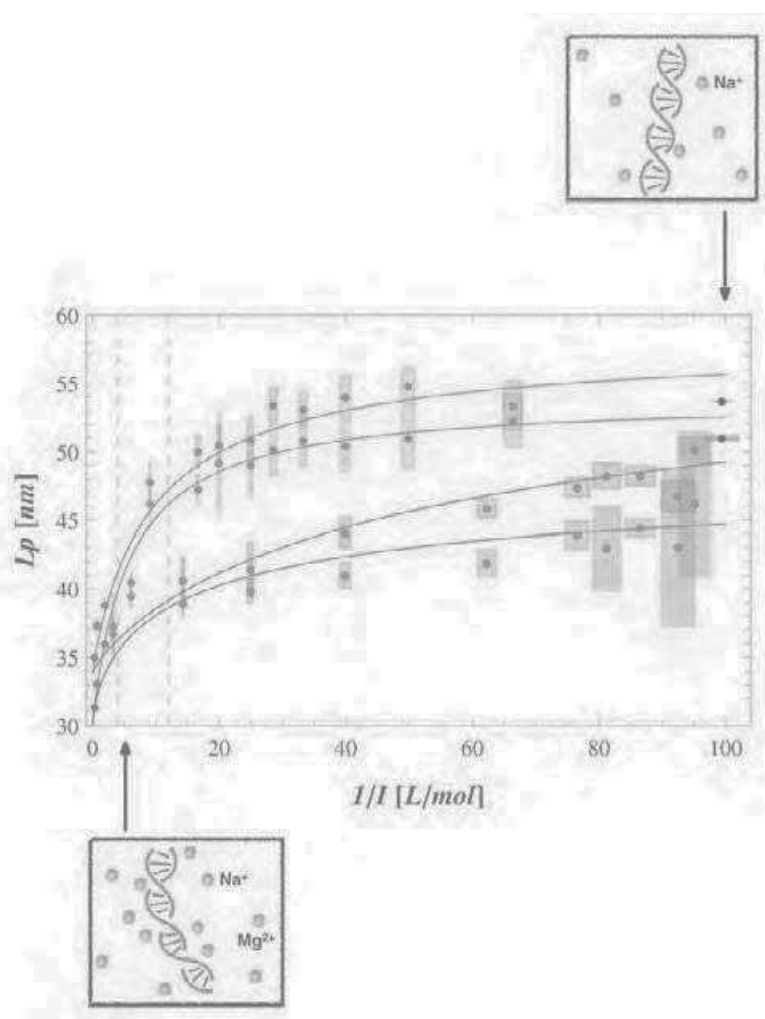
Car, là encore, de nombreuses théories existent dans le domaine afin d'exprimer la variation de  $L_p$  en fonction de la concentration et de la nature des ions présents en solution. L'idée apportée par la théorie OSF permet de quantifier qualitativement la forte variation de  $L_p$  à fort  $I$ . En revanche, cette dernière ne peut être utilisée pour décrire les fluctuations observées aux  $I$  faibles et intermédiaires. Sur cette gamme, la théorie OSF améliorée par la théorie de condensation de Manning décrit bien les fluctuations de  $L_p$  en présence d'ions monovalents sur cette gamme restreinte. Cette théorie ne permet pas la description de  $L_p$  sur toute la gamme de sel explorée. L'ajustement des données en présence d'ions divalents, nécessite de laisser libre la valeur du préfacteur, dont l'évaluation s'éloigne très fortement de la valeur théorique. Ces deux théories, sûrement les plus connues, ne sont donc pas suffisantes pour estimer et évaluer correctement la dépendance de  $L_p$  sur une gamme étendue de force ionique.

L'approche de Netz et Orland en champ moyen propose une description convenable en présence d'ions  $Mg^{2+}$ , mais elle ne permet pas plus d'exprimer la chute aux fortes concentrations en sel observée en présence d'ions monovalents. Seule la dernière théorie de Manning permet de décrire plus convenablement cette brusque chute ainsi que la variation de  $L_p$  sur le reste de la gamme de  $I$  en présence d'ions monovalents. En revanche, cette théorie ne rend pas justement compte de la variation observée en présence d'ions divalents.

Encore aujourd'hui l'influence de la force ionique sur la longueur de persistance n'est pas entièrement comprise. Seule la nouvelle théorie de Manning s'approche d'un accord quantitatif avec l'expérience. En revanche, la prise en compte de l'impact et de l'importance de la nature des ions n'est pas pleinement mise en équation. Ce domaine laisse encore de nombreuses questions en suspens et offre de nombreux champs d'investigations.

Dans cette partie, nous avons aussi pu mettre en exergue l'influence de la séquence même de l'ADN sur ses propriétés physiques. En effet, la présence d'une séquence *Poly(dA)* de 60 bp influe fortement sur le comportement et la réponse du polymère soumis à une variation de concentration en sel. Ces différences sont observables quelle que soit la nature des ions présents en solution. Ces résultats font écho à quelques papiers de la littérature. La mise en lumière de cet effet souligne l'importance de la séquence de l'ADN et de l'implication probable dans différents processus biologiques. Comprendre les rouages de la nature même de cet effet nécessite de plus amples travaux. L'observation de cet effet séquence interroge sur les propriétés et répercussions que peut engendrer une modification structurale locale de l'ADN. Une telle modification, qui peut-être considérée comme un défaut local dans la structure du double brins d'ADN, peut influencer le comportement global de la molécule. En effet, ici, une modification de 60 bp, soit environ d'un dixième de la séquence, influence les propriétés physiques de l'ensemble de la molécule d'ADN de 583 bp. *In vivo*, la fixation de protéines sur des sites spécifiques de l'ADN, allant de 3 à quelques dizaines de paires de base, engendre de fortes modifications de ce dernier. L'existence de tels sites, ou de séquences spécifiques influe-t-elle sur la physique de l'ensemble de la molécule ? La présence de défauts locaux intrinsèques à la molécule d'ADN, autres que la fixation de protéines, est-elle détectable ? L'impact de la présence de ces derniers, telles que des courbures intrinsèques ou des bulles de dénaturation, est-il mesurable ? Le dernier chapitre va nous apporter un certain nombre d'éléments de réponses.

ARTICLE : *Dependence of DNA Persistence Length on Ionic Strength of Solutions with Monovalent and Divalent Salts : A Joint Theory–Experiment Study*



# Dependence of DNA Persistence Length on Ionic Strength of Solutions with Monovalent and Divalent Salts: A Joint Theory–Experiment Study

Annaël Brunet,<sup>†,‡,§</sup> Catherine Tardin,<sup>†,‡</sup> Laurence Salomé,<sup>†,‡</sup> Philippe Rousseau,<sup>||,⊥</sup> Nicolas Destainville,<sup>#,§</sup> and Manoel Manghi<sup>\*,#,§</sup>

<sup>†</sup>CNRS, Institut de Pharmacologie et de Biologie Structurale (IPBS) 205 route de Narbonne, BP 64182, F-31077 Toulouse, France

<sup>‡</sup>UPS, IPBS, Université de Toulouse F-31077 Toulouse, France

<sup>#</sup>UPS, Laboratoire de Physique Théorique (IRSAMC), Université de Toulouse, F-31062 Toulouse, France

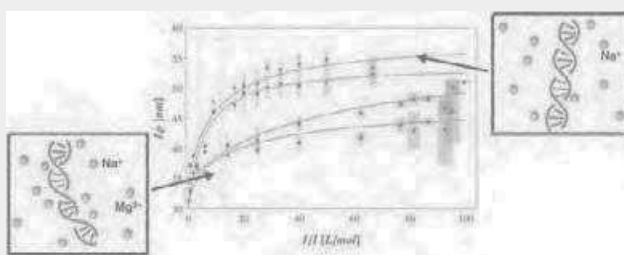
<sup>§</sup>CNRS, Laboratoire de Physique Théorique (IRSAMC), F-31062 Toulouse, France

<sup>||</sup>UPS, Laboratoire de Microbiologie et Génétique Moléculaires (LMGM), Université de Toulouse, F-31062 Toulouse, France

<sup>⊥</sup>CNRS, LMGM, UMR CNRS-UPS 5100, F-31062 Toulouse, France

## S Supporting Information

**ABSTRACT:** Using high-throughput tethered particle motion single-molecule experiments, the double-stranded DNA persistence length,  $L_p$ , is measured in solutions with  $\text{Na}^+$  and  $\text{Mg}^{2+}$  ions of various ionic strengths,  $I$ . Several theoretical equations for  $L_p(I)$  are fitted to the experimental data, but no decisive theory is found which fits all the  $L_p$  values for the two ion valencies. Properly extracted from the particle trajectory using simulations,  $L_p$  varies from 30 to 55 nm, and is comparable to previous experimental results. For the  $\text{Na}^+$ -only case,  $L_p$  is an increasing concave function of  $I^{-1}$ , well fitted by Manning's electrostatic stretching approach, but not by classical Odijk–Skolnick–Fixman theories with or without counterion condensation. With added  $\text{Mg}^{2+}$  ions,  $L_p$  shows a marked decrease at low  $I$ , interpreted as an ion–ion correlation effect, with an almost linear law in  $I^{-1}$ , fitted by a proposed variational approach.



## 1. INTRODUCTION

Ions play a major role in the cell, for example, by modifying the protein activity, inducing a voltage between the intracellular and extracellular matrixes, or controlling the DNA packaging in viral capsids or in the nucleus. Complex mechanisms involving DNA, such as its wrapping around histones or its denaturation, will only be thoroughly understood when the effect of mobile ions on the DNA conformation is elucidated, DNA being one of the most charged biopolymers found in nature.

The first quantitative experimental studies of DNA conformational properties as a function of salt concentration were done in 1978 by Harrington,<sup>1</sup> using flow birefringence (FB) experiments, to measure the DNA radius of gyration in dilute DNA solutions. The DNA radius of gyration is intimately related to the DNA persistence length, namely the correlation length of the tangent–tangent correlation function,

$$\langle \mathbf{t}(s) \cdot \mathbf{t}(0) \rangle = \exp(-s/L_p) \quad (1)$$

where  $\mathbf{t}(s)$  is the unit vector tangent to the chain at the point of curvilinear index  $s$ . The persistence length,  $L_p$ , thus characterizes the chain stiffness at small length scales.

Experimentally, the persistence length has not been measured directly, and the required procedure to extract it

has remained a major issue since these first quantitative measurements.<sup>1</sup> Other optical techniques have been used, such as transient electric birefringence (TEB)<sup>2</sup> or magnetic birefringence (MB),<sup>6</sup> linear dichroism (LD),<sup>5</sup> dynamic light scattering (DLS),<sup>3</sup> and force-stretching by optical tweezers (FOT),<sup>7</sup> to estimate the variation of  $L_p$  as a function of the ionic strength  $I$ . In a recent paper, Savelyev<sup>4</sup> reviewed the available experimental data and showed that they could be divided into two groups based on the distinct behaviors of  $L_p$  found at high ionic strength. Indeed, whereas the first group of experimental data<sup>1,5–7</sup> indicated a slow decrease of  $L_p$  with increasing  $I$ , the second<sup>3,8–12</sup> found a significant decrease.

Hence, no global picture emerges yet from the literature. Many reasons can be put forward, such as the difficulty in estimating accurately the ionic strength in buffers, which is not simply equal to the added salt concentration, or the method of extraction of the persistence length from the experimental observables. Indeed, extraction of the variations of  $L_p$  with  $I$  by using FB, MB, and LD techniques is very sensitive to the

Received: April 9, 2015

Revised: May 19, 2015





optical arrangement between the electric field and the molecular axis, which is related to the tedious evaluation of the magnetic/optical anisotropy of a single base-pair (bp). Moreover, in LD experiments, the mechanism of interaction between nucleic acids and the electric field depends on the polarizability of the ionic cloud surrounding the DNA, which therefore requires an additional modeling. In FOT experiments, the high DNA stretching induced by the force modifies the DNA structure and the organization of its ionic cloud. Finally, in DLS experiments, the DNA hydrodynamic radius is estimated from the measured diffusion coefficient. Inferring the persistence length is not easy, especially due to excluded volume effects.

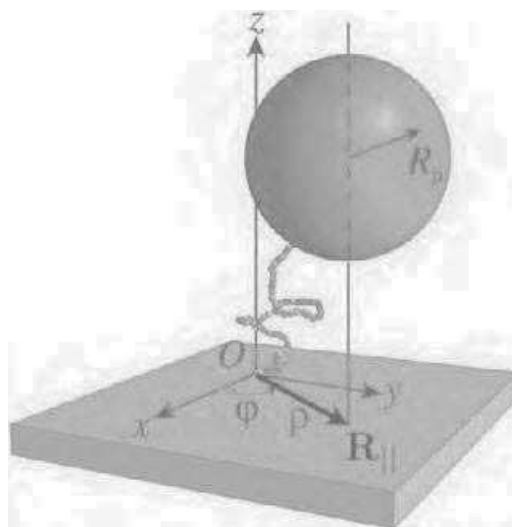
Among the different theories developed to explain the variations of  $L_p$  with the ionic strength,  $I$ , the most famous is the Odijk–Skolnick–Fixman (OSF) theory,<sup>13,14</sup> where  $L_p$  is the sum of a bare, non-electrostatic persistence length and an electrostatic contribution scaling as  $I^{-1}$ . This theory accounts, at least qualitatively, for the fact that a rise in  $I$  provokes an increasing screening of the repulsive phosphate ions of the backbone, which leads to a more flexible DNA chain. Taking into account the so-called Manning counterion condensation around the DNA at low  $I$ , the prefactor in front of  $I^{-1}$ , proportional to the square of the effective DNA charge, is lowered. All the experimental works have therefore been compared to these types of theoretical approaches, with no decisive conclusion.<sup>4</sup> More recently, Manning proposed a new theory,<sup>15</sup> explicitly considering the electrostatic stretching force of the polyelectrolyte, which qualitatively fits some experiments<sup>7</sup> or numerical results at high  $I$ .<sup>4,16</sup>

In this paper we reconsider the old issue, still under debate,<sup>17</sup> of the dependence of DNA conformation, at room temperature, on the ionic strength of the surrounding solution, using the recent single-molecule technique of high-throughput tethered particle motion (HT-TPM).<sup>18,19</sup> We measure the persistence length of two DNAs, of 1201 and 2060 bps, for a large range of well-controlled ionic strengths from  $I = 10^{-2}$  to 3 mol/L with  $\text{Na}^+$  counterions and with or without added  $\text{Mg}^{2+}$  counterions. We assume that the tangent–tangent correlation function is well described by a single persistence length,  $L_p$ , following eq 1. In section 2, we present the simple and well-controlled HT-TPM experiments and their analysis. Section 3 is devoted to the experimental results and the extraction of the persistence length from the HT-TPM amplitude of motion using numerical simulations. Our results are then compared to previous ones in section 4, and to the various existing theories as well as our detailed variational approach in section 5. An interpolation formula that fits all our experimental  $L_p$  values is presented in section 6. Finally, our concluding remarks are given in section 7.

## 2. EXPERIMENTAL SECTION

**2.1. High-Throughput Tethered Particle Motion Experimental Procedure.** DNA molecules were produced by polymerase chain reaction amplification (oligonucleotides from Sigma-Aldrich): Biot-F1201 5'-CTGGTGAGTACTCAACCAAG-3' and Dig-R1201 5'-CTACAATCCATGCCAACC-3' on pTOC1 plasmid, and Biot-F2060 5'-CTGCAATGATACCGCGAGAC-3' and Dig-R2060 5'-TGAC-TTCCGCGTTTCCAGAC-3' on pBR322.

HT-TPM permits the simultaneous tracking of hundreds of single DNA molecules free to fluctuate in solution that are tethered to a coverslip at one end and labeled by a 300 nm particle at the other end (see Figure 1). HT-TPM on-chip assembly is performed as previously described in detail in ref 19. Anchoring of the DNA–particle



**Figure 1.** Principle of the tethered particle motion experiment. The DNA is tethered at one end on the coverslip and labeled by a particle (radius  $R_p$ ) at the other end. The projected particle position  $R_{\parallel}$  is tracked as a function of time. The DNA is modeled by a chain of beads in the simulations.

complexes to the neutravidin (Invitrogen) printed sites was performed in phosphate-buffered saline (Euromedex) supplemented with 1 mg/mL of pluronic 127 (Sigma-Aldrich) and 0.1 mg/mL bovine serum albumin (Sigma-Aldrich), noted T-BSA-Plu.

A large range of buffers were then used to test the effects of both ion valency, by using monovalent  $\text{Na}^+$  or divalent  $\text{Mg}^{2+}$ , and salt concentration on the DNA conformations (Tables I and II in the Supporting Information). The first buffer, corresponding to zero salt added and called zero-salt-buffer in the following, is a phosphate buffer ( $\text{KH}_2\text{PO}_4$  1 mM,  $\text{Na}_2\text{HPO}_4$  3 mM, pH 7.4, pluronic F127 1 mg/mL). To this, we added successively various concentrations of  $\text{NaCl}$  or  $\text{MgCl}_2$  to obtain a large range of salt conditions (X-salt-buffer). Before starting the experiment, the flow cell was extensively rinsed with the zero-salt-buffer ( $\sim 100$  chamber volumes), left to incubate for 1 h at room temperature, and then rinsed again with  $\sim 100$  chamber volumes of zero-salt-buffer. The experiment started with a zero-salt-buffer measurement; next, the concentration of monovalent ions was progressively increased by addition of  $\sim 100$  chamber volumes of X-salt-buffer. The flow cell was extensively rinsed with the zero-salt-buffer ( $\sim 100$  chamber volumes), incubated for 4 min, and rinsed again with  $\sim 100$  chamber volumes of zero-salt-buffer. A new zero-salt-buffer measurement was performed. At last, the divalent ion concentration was progressively increased, and new measurements were carried out. For all conditions, the data acquisition was performed at a controlled temperature of 25 °C, and 1 min movie was recorded and analyzed. We ensure the reliability of the experimental procedure by checking the agreement between the two values of the zero-salt measurement obtained before the addition of monovalent ions and before the addition of divalent ions. Experiments were repeated on different days to ensure the reproducibility of our results.

The tethered particles of 300 nm diameter were visualized using a dark-field microscope (Axiovert 200, Zeiss) equipped with a 32 $\times$  objective, an additional 1.6 $\times$  magnification lens, and a temperature control system (Physitemp TS-4MPER). Images were acquired during 1 min at a frame recording rate of 25 Hz on a Dalsa Falcon 1.4M100 CMOS camera. The field of observation covers an area of 215  $\mu\text{m} \times 160 \mu\text{m}$ .

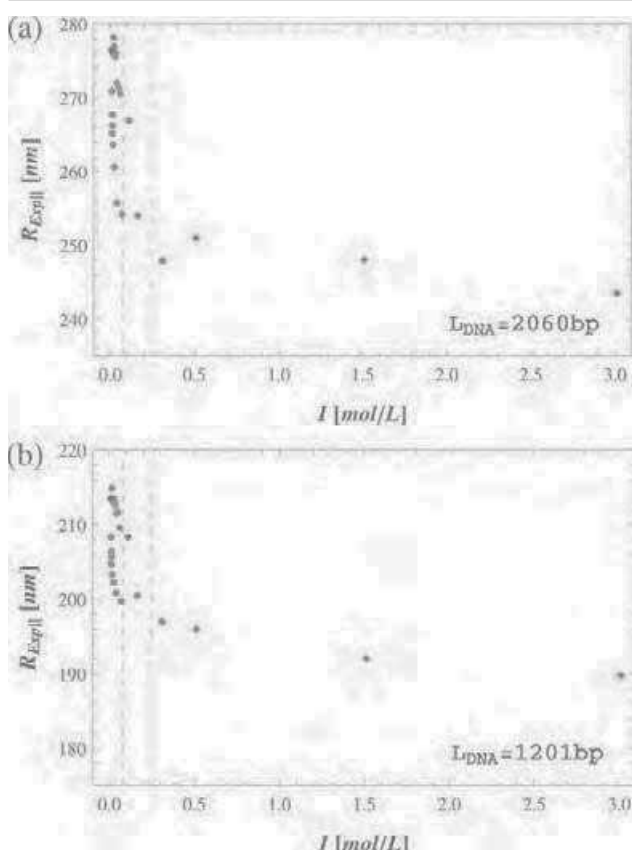
**2.2. HT-TPM Procedure of Analysis.** The software developed by Magellium (France) tracks in real time the positions of all the particles, corrects for experimental drift, calculates the asymmetry factor to select tethered particles valid for the analysis, and finally projects the experimental root-mean-square end-to-end distance on the surface,  $R_{\text{exp}\parallel} \equiv \sqrt{\langle R_{\parallel}^2 \rangle}$ , the amplitude of motion of the bead, along the time

trace (see Figure 1). We refer the reader to ref 21 for the detailed calculations of  $R_{\text{exp||}}$ .

In order to measure the small differences expected on  $R_{\text{exp||}}$ , we set up a two-step procedure for the analysis of  $R_{\text{exp||}}$  where first a criterion of validity and then some corrections are applied. This procedure was performed with home-built Mathematica scripts. Details can be found in ref 19. During this procedure, around 12% of the data was eliminated, and the final number of trajectories kept for each DNA condition typically ranged between 100 and 1000 (see Tables I and II in the Supporting Information).

### 3. EXPERIMENTAL RESULTS

#### 3.1. Simple Estimate of DNA Persistence Length from the Amplitude of Motion. In Figure 2 are shown the



**Figure 2.** Experimental HT-TPM amplitude of motion vs the ionic strength  $I$  for a DNA of length (a)  $L = 2060$  bp and (b)  $L = 1201$  bp with  $\text{Na}^+$  (red) and  $\text{Mg}^{2+}$  (blue) cations. Vertical lines delimit the close-to-physiological salt conditions, as listed in Table 1.

variations of the experimental amplitude of motion, i.e., the root-mean-square end-to-end distance projected on the surface,  $R_{\text{exp||}}$ , measured by HT-TPM, as a function of the ionic strength,  $I$ , of the various buffers listed in Tables I and II in the Supporting Information for two DNA of lengths 2060 and 1201 bp. The ionic strength is defined as

$$I = \frac{1}{2} \sum_i z_i^2 c_i \quad (2)$$

where  $z_i$  and  $c_i$  are respectively the valency and the concentration of ion  $i$  in the buffer. The red (resp. blue) symbols correspond to the monovalent  $\text{Na}^+$  ions (resp. divalent  $\text{Mg}^{2+}$  ions) added to the buffers. The black symbols correspond to zero-salt-buffer conditions.

The plots clearly show that, when  $I$  is increased,  $R_{\text{exp||}}$  decreases by about 10% for monovalent as well as divalent ions. This is thus a subtle effect. This behavior is qualitatively due to the fact that negatively charges carried all along the DNA repel each other more strongly when the ionic strength is low; this repulsion decreases when the ionic strength increases due to electrostatic screening. This self-repulsion increases the rigidity of the macromolecule, thus increasing  $R_{\text{exp||}}$ . Furthermore, one notices that the variation of  $R_{\text{exp||}}$  is amplified when the DNA length,  $L$ , is increased. This is consistent with the fact that the process under examination occurs homogeneously along the DNA.

Comparing the experimental results for monovalent and divalent ions over the same  $I$  range, one observes that the decrease of  $R_{\text{exp||}}(I)$  is more important with  $\text{Mg}^{2+}$  ions, and once again the effect is amplified for long DNAs. Moreover, one notes a sudden drop of  $R_{\text{exp||}}$  at very low concentrations of  $\text{Mg}^{2+}$  (from  $I = 10.1$  mmol/L without  $\text{Mg}^{2+}$  to 10.5 mmol/L). However, it is difficult to know with high precision the ionic strength for such low values, since we cannot exclude that some additional ions are released from the surface, for instance. Therefore, the errors bars are potentially large for these points.

To access the properties of the DNA only, such as its end-to-end distance,  $R_{\text{DNA}} \equiv \sqrt{\langle R_{\text{DNA}}^2 \rangle}$ , it is necessary to correct for the effects of the particle and the glass substrate. The simplest way is to subtract the particle radius by assuming that the DNA extremity and the particle, of radius  $R_p = 150$  nm, move independently, which leads to<sup>19</sup>

$$\langle R_{\text{DNA}}^2 \rangle = \frac{3}{2} \langle R_{\text{||}}^2 \rangle - R_p^2 \quad (3)$$

By doing so, we neglect the effect of the glass substrate. It has been taken into account analytically only for long and flexible polymers,<sup>23</sup> which is not the case for the DNA in this study. Indeed, their length ( $L = 2060$  and 1201 bp, i.e., 700 and 408 nm, using 1 bp = 0.34 nm) is no more than a few persistence lengths ( $L_p \approx 50$  nm), which allows us to qualify them as semiflexible. As already shown in ref 19, the approach proposed in ref 23 does not work well for such DNA lengths. [The approach proposed in ref 23 yields essentially the same result as eq 3 for  $N = 2060$  and is not valid for  $N = 1201$ .]

Figure 1 in the Supporting Information shows the DNA end-to-end distance,  $R_{\text{DNA}}$ , obtained by eq 3, as a function of the inverse of the ionic strength,  $I^{-1}$ . Of course, the relative effect of the salt is slightly higher once the particle radius is deduced. The next step is to extract the DNA persistence length,  $L_p$ , from  $R_{\text{DNA}}$ . The simplest way would be to use the wormlike chain (WLC) formula, valid for a phantom chain in solution, resulting from eq 1:<sup>24</sup>

$$\langle R_{\text{DNA}}^2 \rangle = 2L_p^2 \left( \frac{L}{L_p} - 1 + e^{-L/L_p} \right) \quad (4)$$

However, this way of extracting  $L_p$  leads to quite high values of the persistence length compared to the commonly accepted values around 50 nm. For instance, at low  $I$ ,  $L_p$  saturates around 76 nm for the 2060 bp long DNA and around 68 nm for the 1201 bp long DNA (data not shown). These high values might be due to the particle–substrate interaction, or particle–polymer or monomer–monomer excluded volume interactions, which may swell the DNA.

**3.2. Refined Extraction of the DNA Persistence Length Using Simulations of the HT-TPM Setup.** To check these

effects and the approximations used in eqs 3 and 4, we performed numerical simulations. DNA–particle conformations are computed numerically by exact sampling.<sup>23</sup> The labeled DNA polymer is generated as a random walk of  $N$  steps, corresponding to the links of length  $2a$ , with a bending energy by step,  $E_{\text{bend}} = -\kappa_b \cos \theta$ , where  $\kappa_b$  is the bending modulus and  $\theta$  is the angle between successive steps. The number of steps  $N$  was chosen such that  $a = 6$  bp, corresponding to the DNA diameter. The starting point is on the substrate, and at each step, self-intersecting trajectories (resp. trajectories intersecting the substrate) are eliminated to take into account intrachain excluded volume interactions (resp. repulsive interactions with the substrate). Hence, the polymer is modeled by a chain of beads of (excluded) volume  $v = 4\pi a^3/3 \simeq 36$  nm<sup>3</sup>, which is taken to be constant as a function of the ionic strength (see Figure 1). The salt effects are therefore supposed to be completely taken into account in the bending modulus,  $\kappa_b$ . The last step, of length  $a + R_p$ , has a uniformly random orientation. The persistence length value is related to  $\kappa_b$  by  $L_p = 2a\beta\kappa_b$ , where  $\beta = (k_B T)^{-1}$  is the inverse of the thermal energy. [Strictly speaking, the discrete WLC persistence length leads to  $L_p = -2a/\ln[\coth(\beta\kappa_b) - 1/\beta\kappa_b]$ . For  $L_p \in [35, 70]$  nm, the error is less than 0.25 nm when using the approximation  $L_p \simeq 2a\beta\kappa_b$ .] The two-dimensional projection of the particle position,  $R_{||}$ , was measured, and the amplitude of motion, defined as  $R_{||} \equiv \sqrt{\langle R_{||}^2 \rangle}$ , is averaged over several million independent trajectories. Since at this level of coarse-graining, electrostatic interactions are not included, we varied  $\kappa_b$  by hand such that  $L_p$  spans the range 36–70 nm, in order to reflect the stiffening due to the decrease of the ionic strength in the solution. We do not consider the torsional degrees of freedom in these simulations, since no constant torque is applied on the particle and therefore on the DNA. Hence, the end-to-end distance (eq 4), the quantity of interest here, is not modified by torsional fluctuations as soon as, at this level of coarse-graining, the DNA is locally viewed as a straight cylinder (model KP1 as defined in ref 25).

Figure 3 shows the simulated  $R_{||}$  as a function of  $L_p$  (full circles). We observe an increase of  $R_{||}$  from 240 to 300 nm for the 2060 bp long DNA, and from 195 to 230 nm for the 1201 bp long DNA. Both ranges of  $R_{||}$  values contain the corresponding experimental observations, which thus indicates that we explored a good range of persistence length values.

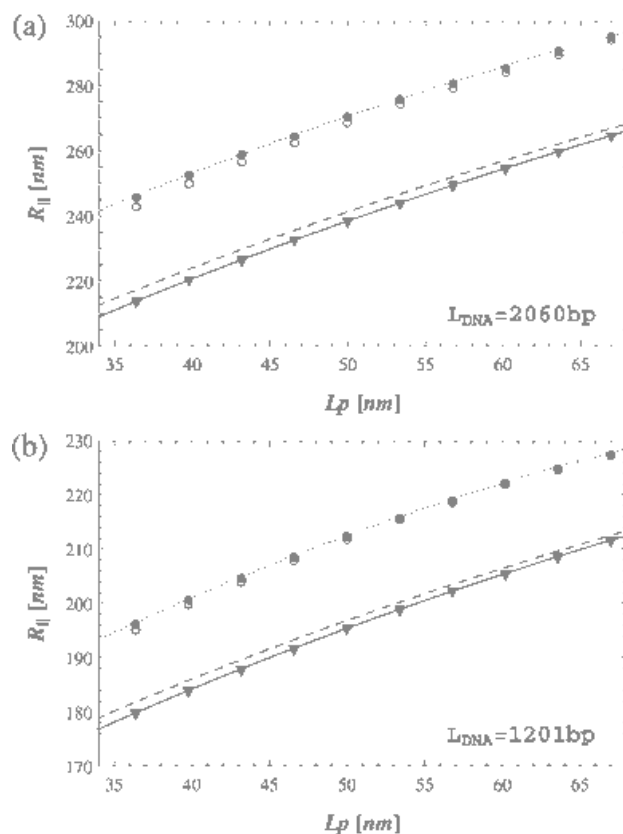
For the purpose of comparison, we also plotted the end-to-end distance of a free polymer–particle complex without any excluded volume interactions and without wall (triangles), and the polymer–particle plus substrate without excluded volume (open circles). The solid curves correspond to the discrete WLC result for the end-to-end distance, using eq 3 (to include the particle contribution), and without any excluded volume and wall:<sup>26</sup>

$$\langle R_{||}^2 \rangle = \frac{2}{3} [a^2 N W_N(v(\kappa_b)) + R_p^2] \quad (5)$$

where

$$W_N(x) = \frac{1+x}{1-x} - \frac{2x}{N} \frac{1-x^N}{(1-x)^2} \quad (6)$$

and  $v(\kappa_b) = \coth(\kappa_b) - 1/\kappa_b$ . Clearly, the solid curve perfectly matches the simulation results, as expected. The dashed line corresponds to the continuous WLC (no excluded volume and no wall), eq 4, which gives slightly larger end-to-end distances.

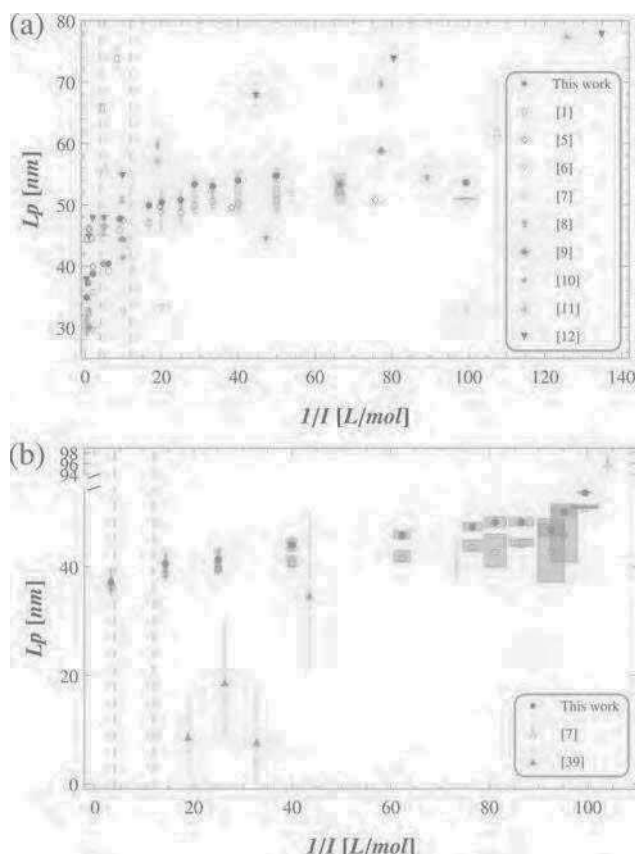


**Figure 3.** Simulated  $R_{||}$  vs  $L_p$  for DNA of (a) 2060 bp ( $N = 172$ ) and (b) 1201 bp ( $N = 100$ ). Triangles correspond to a free polymer–particle complex without excluded volume. Full (resp. open) circles correspond to a grafted polymer in the TPM geometry with a hard core particle and with (resp. without) intrachain excluded volume interactions (see text). The solid line is the discrete WLC formula, eq 5; the dashed line is the continuous WLC formula, eq 4; and the dotted lines are fits by polynomials: (a)  $R_{||} = (-10.7 \times 10^{-3})L_p^2 + 2.71L_p + 162.1$  [nm], and (b)  $R_{||} = (-9.0 \times 10^{-3})L_p^2 + 1.955L_p + 137.35$  [nm].

One observes that the presence of the particle that interacts with both the substrate and the chain induces a non-constant shift to higher values of  $R_{||}$  (from triangles to full circles). Since the intrachain excluded volume swells the polymer by less than 2 nm, especially for small values of  $L_p$  (more flexible chains), the main difference comes from the substrate–particle interactions. This is why the extraction of  $L_p$  using eqs 3 and 4, or equivalently eq 5, overestimates  $L_p$  by about 20 nm.

To obtain precise values of  $L_p$  from the experiments, we thus fitted the simulation data  $R_{||}(L_p)$  by a quadratic polynomial law (see Figure 3), which in turn allows us to accurately determine the experimental  $L_p$  from the experimental values of  $R_{||}$ . The persistence length is then plotted as a function of  $I^{-1}$  in Figure 4 for the two DNA lengths and the two types of counterions (figures are given in Tables I and II in the Supporting Information). Other available data found in the literature are also shown.

With Na<sup>+</sup> counterions,  $L_p$  values are in the same range for both  $L = 2060$  and 1201 bp, which tends to confirm that the persistence length extracted with this procedure is almost independent of the DNA length, as expected. It increases monotonically from roughly 35 nm for high ionic strength ( $I \simeq 3$  mol/L) to 54 nm for low ionic strength ( $I = 10$  mmol/L),



**Figure 4.** DNA persistence length,  $L_p$ , vs the inverse of the ionic strength,  $\Gamma^{-1}$ , extracted from the HT-TPM data (full circles for  $L = 2060$  bp and open circles for  $L = 1201$  bp), and from available literature data (references are given in the legend). Buffers with (a) monovalent  $\text{Na}^+$  salt counterions and (b) added divalent  $\text{Mg}^{2+}$  counterions.

which corresponds to an increase of more than 50%. Near physiological salt conditions, around 150 mmol/L, we find  $L_p \approx 43$  nm. Moreover, using this plot representation, the data show a clear concave shape.

With  $\text{Mg}^{2+}$  counterions,  $L_p$  is greatly reduced, which is a signature of the role of the ion valency  $z$ . This has already been observed in previous experiments.<sup>2,7</sup> The  $L_p$  values, between 35 and 50 nm, are slightly different for the two DNAs,  $L_p$  being larger by almost 5 nm at low  $I$  for the longest DNA. Moreover, we observe an abrupt decrease of  $L_p$  between the case of no divalent ions and the previous  $\Gamma^{-1}$  value, corresponding to the addition of 0.15 mmol/L of  $\text{Mg}^{2+}$ . At higher  $I$ , the increase of  $L_p$  is almost linear in  $\Gamma^{-1}$ .

In the following, we shall try to fit the so-obtained DNA persistence lengths using the available theories found in the literature. Before this, we compare our experimental values to the ones found in the literature.

#### 4. COMPARISON WITH OTHER EXPERIMENTAL DATA

**4.1. Influence of  $\text{Na}^+$  Monovalent Ions.** We start the comparison with other experimental studies by focusing on the measurements of  $L_p$  for values of  $I$  close to the physiological salt conditions. The behavior of  $L_p$  under the action of ionic strength is then discussed.

*Persistence Length Close to the Physiological Salt Concentration.* An inventory of the values of  $L_p$  measured

near physiological salt conditions, i.e., for  $I \in [100, 200]$  mmol/L with only monovalent  $\text{Na}^+$  counterions, is presented in Table 1. The mean value of these  $L_p$  values found in refs 5, 7–12, and

**Table 1.** Summary of  $L_p$  Measured in the Close-to-Physiological Salt Conditions (with  $\text{Na}^+$ ) Found in the Literature<sup>a</sup>

$I$ [mmol/L]	$L_p$ [nm]	$L_{\text{DNA}}$ [bp]	experimental method	ref
223.0	$66 \pm 3$	6646	FB (25 °C)	1
210.8	46.5	6646	DLS (20 °C)	8
210.8	54.7	6646	DLS (20 °C)	11
210.8	40.6	6646	DLS (20 °C)	9
204.1	$48 \pm 2$	39936	DLS (20 °C)	12
201.0	$40.6 \pm 0.4$	39936	FB (25 °C)	10
201.0	$46.8 \pm 0.4$	39936	LD	5
186.2	$56 \pm 3$	48502	FOT (25 °C)	7
165.1	$40.5 \pm 0.4$	2060	HT-TPM (25 °C)	this work
165.1	$39.5 \pm 0.5$	1201	HT-TPM (25 °C)	this work
154.0	$50 \pm 5$	434, 587	TEB (20 °C)	2
123.0	$74 \pm 3$	6646	FB (25 °C)	1
110.1	$47.8 \pm 0.7$	2060	HT-TPM (25 °C)	this work
110.1	$46.2 \pm 0.8$	1201	HT-TPM (25 °C)	this work
103.1	44.6	6646	DLS (20 °C)	8
103.1	44.6	6646	DLS (20 °C)	9
103.1	$53 \pm 2$	39936	DLS (20 °C)	12
102.4	$55 \pm 2$	39936	DLS (20 °C)	12
101.1	$43 \pm 1$	43–179	TED (20 °C)	28
101.0	$47.8 \pm 0.4$	39936	LD	5
101.0	$41.8 \pm 0.6$	39936	FB (25 °C)	10
93.4	$43 \pm 1$	48502	FOT (25 °C)	7

<sup>a</sup>In order to compare the whole set of published data,<sup>5,7–12,28</sup> we rigorously computed the ionic strength used in these data by directly taking experimental values when available, or the values deduced by interpolation otherwise.

28 is  $48 \pm 6$  nm, which is in good agreement with our interpolated value:  $L_p(150 \text{ mmol/L}) = 43 \pm 3$  nm. In fact, the measured  $L_p$  varies widely with the experimental techniques used, with the studied DNA (with lengths varying from 6646 bp to 50 kbp), and with the theoretical and analytical tools used to extract  $L_p$ . In addition, the variability of these  $L_p$  values might be attributed to the difficulty in perfectly controlling the presence of divalent ions such as  $\text{Mg}^{2+}$ , the presence of which can have a dramatic effect even at low concentrations (on the order of mmol/L), as mentioned before. Finally, it can be noted that the commonly accepted value of  $L_p = 50$  nm for a “random” DNA sequence at physiological salt conditions, i.e.,  $I \approx 150$  mmol/L ( $\Gamma^{-1} \approx 6.7$  L/mol) with only monovalent  $\text{Na}^+$  counterions,<sup>27</sup> slightly exceeds the experimentally derived values.

*Variation of  $L_p$  on the Whole  $I$  Range.* In order to compare the global behavior of  $L_p(I)$  that we measured with the previously published results, we superimposed all the results in Figure 4a, where  $L_p$  values are plotted as a function of  $\Gamma^{-1}$ . As mentioned in the Introduction, Savelyev<sup>4</sup> recently reviewed all the available experimental data and showed that they could be divided into two groups, based on the difference in  $L_p$  behaviors observed at high ionic strength,  $0.11 \leq I \leq 3$  mol/L (see Figure 4). We will keep this division to compare our results to those obtained from the first group of experimental data,<sup>1,5–7</sup> which shows a slow decrease of  $L_p$  with increasing  $I$ , and then to those

obtained from the second group of data,<sup>3,8–12</sup> which found a significant decrease of  $L_p$ .

In the first set of experiments, the researchers observed a saturation of  $L_p$  at high  $I$ . More precisely, for  $I$  exceeding 150 mmol/L,  $L_p$  decreased by only 10%. Moreover, at low  $I$ , these publications showed an increase in  $L_p$  of about 10%, similar to our observations from our HT-TPM measurements. It can be noted that Harrington<sup>1</sup> found a larger increase of  $L_p$  in the low  $I$  range than all the other results.

This first group of experimental works gathers results obtained by FB from Harrington,<sup>1</sup> MB from Maret et al.,<sup>6</sup> FOT from Baumann et al.,<sup>7</sup> and LD from Rizzo et al.<sup>5</sup> [We recalculated the  $L_p$  values (between 30 and 90 nm) reported Maret et al. by using the raw data collected in Table 1 and eq (5) of ref 6. Our calculation does not correspond to the data reported by Savelyev et al.<sup>4,16</sup>] The FB, MB, or LD methods may be prone to perturbative Joule heating and bulk electrophoresis effects; to minimize these effects, the ionic strength was kept low. As the FOT method induces a perturbation of the sample structure in the high force regime, we consider only the  $L_p$  values obtained in the low force regime, using the inextensible WLC model as a comparison to our HT-TPM results, where no force is applied. It is important to note that these studies were performed on DNA exceeding 40 kbp long, and that only scarce measurements were performed at high ionic strength, in opposition to the second group of data.

In this regime of high ionic strength, the second group did not show a plateau but rather a significant decrease of  $L_p$  by about 25–30%. This observation is in good agreement with our measured decrease of 25% in the same  $I$  range. In the low ionic strength range,  $0.01 \leq I \leq 0.1$  L/mol, this second group measures a small but regular decrease of  $L_p$  by about 15% when  $I$  increases, which is larger than the decrease we measured, about 10%. FB data<sup>10</sup> can be classified in this set of data, given the 25% increase of  $L_p$  in the high  $I$  range. However, variations in the low ionic strength range seem to be modest in comparison to the other publications of this set, but closer to ours.

This second group of experimental data is essentially based on the DLS experiments performed by Sobel et al.<sup>12</sup> and Kam et al.<sup>9</sup> Manning<sup>8</sup> and Post<sup>11</sup> proposed corrections to the extracted  $L_p$  values from the original data reported by Borochoy et al.<sup>3</sup> In those DLS experiments, the DNA hydrodynamic radius was deduced from the diffusion coefficient measurement. To infer  $L_p$ , the usual Gaussian polymer model was used. The number of Kuhn segments being large for the 6646 bp (resp. ~40 kbp) DNA under study,  $N \simeq 22$  (resp.  $N \simeq 1333$ ), the swelling of the chain was induced as a result of excluded volume. Therefore, it is essential to precisely estimate the excluded volume. This is nevertheless a challenging task. For instance, Manning's<sup>8</sup> and Post's<sup>11</sup> corrections led to  $L_p$  values differing by 4 nm at  $I = 8$  mmol/L and by ~1 nm at  $I = 1$  mol/L.

Note that we compare studies performed on long DNA, from 6646 bp to 40 kbp, which are in the flexible regime and thus much more sensitive to excluded volume effects than our HT-TPM experiments made on DNA of lengths  $L = 2060$  or 1201 bp, which are in the semiflexible regime.

In this quantitative comparison, we only considered experiments performed on the same range of ionic strength induced by  $\text{Na}^+$  ions. As a result, refs 2 and 27–34, which studied the effect of very low  $\text{Na}^+$ -induced  $I$ , as well as refs 35

and 36, which focused on the influence of  $\text{K}^+$  or  $\text{Li}^+$  ions, and ref 37, which studied the combined effect of multivalent and monovalent ions, were discarded from our comparison. Note that the single-molecule study in ref 38 was not considered either because too few values of  $I$  were explored.

**4.2. Influence of  $\text{Mg}^{2+}$  Divalent Ions.** The decrease of  $L_p$  that is measured when  $I$  is varied from 10 to 310 mmol/L is about 35% whatever the ion valency and the DNA length (Figure 4). Yet the shape of this decrease as a function of  $I$  is completely different, depending on the valency of the ions used. Divalent ions appears to induce an almost linear decrease of  $L_p$  as a function of  $I^{-1}$ , while monovalent ions cause a decrease with a concave shape. In addition, the absolute value for  $L_p$  is smaller by about 5 nm with divalent ions than with monovalent ions at low  $I$ . This observed trend is in agreement with the first quantitative observations at very low  $I$  by Hagerman,<sup>2</sup> Elias and Eden,<sup>29</sup> and then others.<sup>6,7</sup> In addition, Hagerman<sup>2</sup> observed on 434 and 587 bp DNA by TEB the same abrupt decrease of  $L_p$  at very low  $I$  with  $\text{Mg}^{2+}$  ions, followed by a slower decrease, as shown in Figure 7. The magnitude of this initial decrease was, however, larger by around 30% in ref 2, and 60% in ref 7, whereas ref 6 found a decrease similar to ours.

Some other studies explored the influence of  $\text{Mg}^{2+}$  on DNA flexibility<sup>28,32,34,36,40,41</sup> and showed a rise in DNA flexibility with the addition of  $\text{Mg}^{2+}$  counterions in solution. Nevertheless, these studies probed only a few values of  $I$ , and some KCl was added to the solution, preventing any quantitative comparison.

Dietrich et al.<sup>39</sup> also used the TPM technique to monitor the effect of  $\text{Mg}^{2+}$  on a DNA fragment of 4882 bp and observed a large decrease of  $R_{\text{exp}}$  in the presence of divalent ions. Yet the  $L_p$  values extracted from these experimental data, much smaller than any other published experimental data, cannot be quantitatively compared to our results due to several errors in the extraction procedure. The persistence length was extracted assuming that the particle excursion was related to  $L$  using a simple Hooke law in the Gaussian regime,  $\langle R^2 \rangle = 2L_p L$ , thus seemingly forgetting the factor 3/2 due to dimensionality and not subtracting the particle radius (see eq 3), but also ignoring excluded volume effects. Moreover, they presented a very large error bar of 10 nm, due to the small number of trajectories, ranging between 6 and 27.

In the next section, we compare our experimental values of the DNA persistence length to the various theories developed in the literature.

## 5. COMPARISON BETWEEN EXPERIMENTS AND EXISTING THEORIES

**5.1. Odijk–Skolnick–Fixman Approach at High Ionic Strength.** Several models have been proposed in the literature to explain the variation of the persistence length,  $L_p$ , of polyelectrolytes with the ionic strength,  $I$ . When electrostatic interactions between mobile ions and the polyelectrolyte are taken into account at the Debye–Hückel level (mean-field level and approximation of small values of the electrostatic potential), it has been shown by Odijk<sup>13</sup> and Skolnick and Fixman<sup>14</sup> (OSF) that, using a perturbative approach around an infinitely stiff rod, the persistence length has two contributions,

$$L_p^{\text{OSF}} = L_p^\infty + \frac{l_B}{4A^2\kappa^2} \quad (7)$$

where  $L_p^\infty$  is the bare persistence length (in the limit  $\kappa A \rightarrow \infty$ ), and the second term is an electrostatic contribution to  $L_p$ , where  $A$  is the distance between elementary charges along the chain and

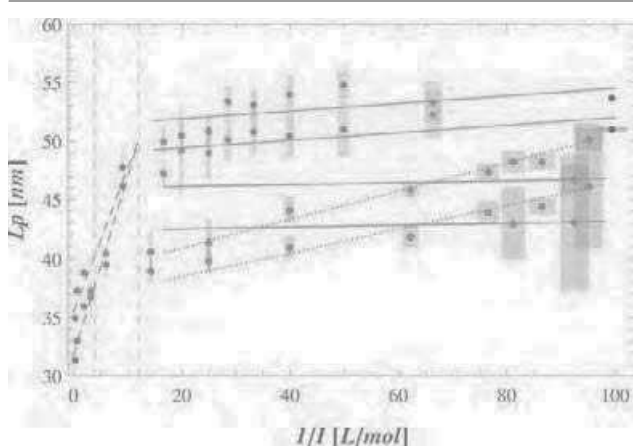
$$\kappa = \sqrt{8\pi l_B I} \quad (8)$$

is the Debye–Hückel screening parameter. The Bjerrum length,  $l_B = e^2/4\pi\epsilon k_B T$ , is equal to 0.715 nm in water at room temperature, which yields  $\kappa = 3.29\sqrt{I} \text{ nm}^{-1}$  (where  $I$  is in mol/L). Equation 7 is theoretically valid for polymer conformations close to the rodlike one, i.e., for  $l_B L_p^\infty \gg A^2$ .

For a dsDNA, with two phosphate anions per bp, we have  $A = 0.17 \text{ nm}$  and  $L_p^\infty \approx 50 \text{ nm}$ , so the validity of eq 7 is well verified. As a function of the ionic strength,  $I$ , eq 7 can be rewritten as

$$L_p^{\text{OSF}} = L_p^\infty + \frac{0.559}{I} [\text{nm}] \quad (9)$$

where  $I$  is in mol/L. Note that the numerical value of the coefficient depends on the model, as shown by Fixman.<sup>42</sup> In any case, Fixman affirms that the exponent in  $I^{-1}$  is robust at large  $I$  (large  $\kappa$ ). Clearly the data in Figure 5, where the



**Figure 5.** Linear fits of the persistence length,  $L_p$ , measured by HT-TPM (already shown in Figure 4) vs the inverse of the ionic strength,  $I^{-1}$ . Dashed lines correspond to eq 7 for large  $I^{-1}$ ,  $L_p^\infty + C/I$ , with two fitting parameters (see text for values), and solid lines to eq 11 for low  $I^{-1}$ , with  $L_p^0$  as fitting parameter. Better fits are obtained for  $\text{Mg}^{2+}$  by leaving  $B$  free (dotted lines,  $L_p^0 = 38.5 \text{ nm}$ ,  $B = 0.121$  for  $L = 2060 \text{ bp}$ , and  $L_p^0 = 36.5 \text{ nm}$ ,  $B = 0.1$  for  $L = 1201 \text{ bp}$ ).

experimental  $L_p$  is plotted vs  $I^{-1}$ , show a concave shape for high ionic strength, and therefore are not well fitted by the linear law, eq 9. Figure 5 shows a linear fit (dashed lines),  $L_p^\infty + C/I$ , for  $0 < I^{-1} < 10 \text{ L/mol}$ , which yields  $C = 1.21 \text{ nm}\cdot\text{mol/L}$  and  $L_p^\infty = 35.5 \text{ nm}$  for  $L = 2060 \text{ bp}$ , and  $C = 1.52 \text{ nm}\cdot\text{mol/L}$  and  $L_p^\infty = 31.7 \text{ nm}$  for  $L = 1201 \text{ bp}$ , which are almost 3 times larger than the values predicted by OSF. Since we do not have enough data at very high  $I$  with added  $\text{Mg}^{2+}$ , we did not try to fit eq 9 for the  $\text{Mg}^{2+}$  case.

The data obtained by the first group<sup>1,5–7</sup> at high  $I$  were qualitatively in agreement with the OSF theory, in the sense that the variations of  $L_p$  are small at high  $I$ . However, no fit using eq 7 was done in those works.

**5.2. Manning Charge Renormalization at Low Ionic Strength.** The OSF theory fails to explain the variation of  $L_p(I^{-1})$  for the whole range of  $I$  studied; indeed, it does not

reproduce the concave shape observed in Figure 7. A plausible explanation is that the Debye–Hückel approximation of small values of the electrostatic potential is not valid for double-stranded (ds) DNA, which is highly charged. Indeed, Manning showed in the 1960s that, in the limit of low  $I$ , some counterions somehow “condense” close to the DNA due the large net charge of DNA.<sup>43</sup> This effect is a nonlinear effect associated with the mean-field Poisson–Boltzmann equation close to charged cylinders. This phenomenon, known as the Manning condensation, tends to reduce the effective charge of the DNA. Manning proposed that, in the *low salt limit*, the effective linear charge density is  $\alpha/A$ , where  $\alpha$  depends on the parameter

$$u = \frac{z l_B}{A} \quad (10)$$

where  $z$  is the counterion valency. If  $u < 1$ ,  $\alpha = 1$ , and if  $u > 1$  then  $\alpha = 1/u$ . For dsDNA, the Manning parameter is  $u = 4.21$  for  $\text{Na}^+$  counterions ( $z = 1$ ) and  $u = 8.41$  for  $\text{Mg}^{2+}$  ( $z = 2$ ). Thus, for monovalent counterions,  $\alpha = 0.24$ , and the effective charge is decreased by a factor of roughly 75%. For a mixture of counterions with different valencies, which is the case in our experiments with  $\text{Mg}^{2+}$  ions, the effective charge is  $\alpha = A/zl_B$ , where  $z$  is the largest valency; i.e., only counterions with the largest valency (here divalent ones) condense along the DNA, and no monovalent ions are condensed (unless divalent counterions are depleted, which is not the case in most of the experiments and especially not in ours).<sup>44</sup>

The OSF equation, eq 7, has thus been modified according to<sup>45,46</sup>

$$L_p^{\text{OSFM}} = L_p^0 + \frac{1}{4z^2 l_B \kappa^2} \quad (11)$$

which changes the slope  $B$  of the linear relationship  $L_p = L_p^0 + B/I$  from 0.559 (eq 9) to 0.033 nm·mol/L, and should be valid at low ionic strength, i.e., for large  $I^{-1}$ . This is the reason why the constant  $L_p^0$  is *a priori* different from that in eq 7,  $L_p^\infty$ . Note that adopting the Manning condensation, valid at low  $I$ , to the OSF calculation of the persistence length, valid at high  $I$ , is somewhat inconsistent. Moreover, eq 11 is not a proper asymptotic expansion in the limit  $\kappa A \rightarrow 0$ , since it leads to a diverging  $L_p$ . Of course, this limit cannot be reached in practice, since DNA counterions and ions resulting from water dissociation are always present even when no salt is added, which ensures that  $I \neq 0$  (even if it can be very small). Theoretically, the salt-free case corresponds to  $\kappa^{-1} \gg L$ , i.e., to ionic strengths  $I \ll 10^{-6} \text{ mol/L}$ , which is far from being the case in our experiments.

We fitted eq 11 to our experimental  $L_p$  values at low salt for the  $\text{Na}^+$  case, keeping  $L_p^0$  as a free parameter. The results, shown in Figure 5 (solid lines), are quite satisfactory, with  $L_p^0 = 51.3 \text{ nm}$  for  $L = 2060 \text{ bp}$  and  $48.8 \text{ nm}$  for  $L = 1201 \text{ bp}$  (keeping  $B$  free leads to a slightly higher value of  $B = 0.038 \text{ nm}\cdot\text{mol/L}$ ). However, eq 11 does not fit the  $L_p$  values for the  $\text{Mg}^{2+}$  case, the slope being larger (we found  $B \approx 0.1 \text{ nm}\cdot\text{mol/L}$ ), whereas eq 11 predicts a slope divided by  $z^2 = 4$ ,  $B = 0.008 \text{ nm}\cdot\text{mol/L}$ .

The persistence lengths measured in refs 5–7 and 47, at low  $I$  with  $\text{Na}^+$  ions, were fitted by  $L_p = L_p^0 + B/I$ . In ref 7,  $B = 0.033 \text{ nm}\cdot\text{mol/L}$  was fixed to the Manning value, and  $L_p^0$  was found to be around 45–50 nm. However, as already observed and discussed by Manning,<sup>15</sup> the fit is poor, and the error bars are quite large for low  $I$ . We checked that a parameter value of  $B =$

0.089 nm·mol/L, which does not correspond to any theoretical value, leads to a better fit (not shown). Maret et al.<sup>6</sup> fixed  $L_p^0 = 50$  nm and obtained reasonable fits with  $0.024 < B < 0.041$  nm·mol/L, which suggests a large error bar on the experimental values. Rizzo et al.<sup>3</sup> also fitted their data for  $3 < I < 1000$  mmol/L. They obtained  $L_p^0 = 46 \pm 1$  nm, in agreement with our value (see Figure 5), and  $B = 0.043$  nm·mol/L. Wenner et al.<sup>47</sup> measured  $L_p$  of dsDNA for various  $\text{Na}^+$  concentrations by fitting force–extension curves at low forces. They obtained  $L_p^0 = 46$  nm and  $B = 0.037$  nm·mol/L. All these values of  $B$  are in agreement with ours and the value predicted by eq 11, 0.033 nm·mol/L.

Tomić et al.<sup>48</sup> did dielectric spectroscopy experiments on semidilute DNA solutions with NaCl to investigate the high-frequency and low-frequency relaxation modes vs added salt concentration strength. In the high added salt limit (and relatively low DNA concentration), the length scale of the low-frequency relaxation mode,  $L_{LF}$ , can be interpreted as the DNA persistence length  $L_p$ . Their results are in qualitative agreement with the OSF–Manning theory, but with a coefficient  $B = 0.08$  nm·mol/L larger than the Manning value and smaller than the OSF value. Note that this discrepancy can be due to the fact that, in these experiments, the total ionic strength is different from the added salt concentration. Using the same experimental method, Tomić et al. investigated the effect of  $\text{Mg}^{2+}$  in solution.<sup>49</sup> The  $L_{LF}$  was about 1.5 times shorter in Mg–DNA solution than in Na–DNA solution, also suggesting an increased screening with  $\text{Mg}^{2+}$ . The behavior of  $L_{LF}$  was again explained by the OSF–Manning theory, but with a different value of the effective linear density.

Later, Manning proposed to modify eq 11 by multiplying the salt-dependent persistence length by a factor  $(2u - 1)/u \simeq 1.76$  for  $z = 1$  and 1.88 for  $z = 2$ , which gives a worse result for  $\text{Na}^+$  and is not sufficient for  $\text{Mg}^{2+}$  in our case.<sup>50</sup> This correction is therefore not suitable.

**5.3. Mean-Field Nonlinear Corrections at Intermediate Ionic Strengths.** In any case, eq 11 does not explain the concave shape shown in Figure 7 at intermediate ionic strengths. To do so requires renormalization factor for the DNA charge,  $\alpha$ , which depends on the ionic strength  $I$ . Such an approach was developed by Netz and Orland,<sup>51</sup> in which the Poisson–Boltzmann equation is variationally approximated by a Debye–Hückel equation with  $\alpha$  as a variational parameter which renormalizes the electrostatic potential at the DNA surface.

Here, we do the calculations by assuming that the DNA is not penetrable by ions, contrary to ref 51. The dimensionless electrostatic Debye–Hückel potential,  $\phi = ze\beta\psi$ , for an electrolyte (valency  $z$ ) of Debye–Hückel constant  $\kappa$  around a cylinder of radius  $R$  and surface charge density  $\sigma = 1/2\pi AR$  empty of ions is

$$\phi(r) = \frac{2u}{\kappa R} \frac{K_0(\kappa r)}{K_1(\kappa R)} \quad \text{for } r > R \quad (12)$$

$$\phi(r) = \frac{2u}{\kappa R} \frac{K_0(\kappa R)}{K_1(\kappa R)} \quad \text{for } r \leq R \quad (13)$$

where  $K_0$  and  $K_1$  are the modified Bessel function of order 0 and 1 and  $u = z l_B/A$  is the Manning parameter.

Following Netz and Orland,<sup>51</sup> the full nonlinear Poisson–Boltzmann equation is solved variationally by assuming that the solution is of the Debye–Hückel form,  $\alpha\phi(r)$ , where  $\phi(r)$  is

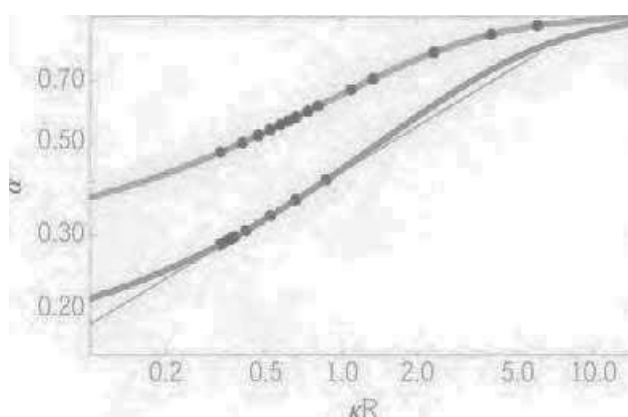
given in eq 12, and  $\alpha$ , the fraction of the “free” counterions, is the variational parameter and is the solution of (see the Supporting Information) the following:

$$4\pi z l_B (1 - \alpha) \int d\mathbf{r} \rho(\mathbf{r}) \phi(\mathbf{r}) = \kappa^2 \int d\mathbf{r} \Omega(\mathbf{r}) \phi(\mathbf{r}) \{ \sinh[\alpha\phi(\mathbf{r})] - \alpha\phi(\mathbf{r}) \} \quad (14)$$

where  $\rho(\mathbf{r}) = \sigma\delta(r - R)$  is the charge distribution. Note that the ionic exclusion factor  $\Omega(\mathbf{r})$  was incorrectly put just in front of the sinh in ref 51. Using eq 12, eq 14 simplifies to

$$2u(1 - \alpha)K_0(\kappa R) = \int_{\kappa R}^{\infty} x K_0(x) \left\{ \sinh \left[ \frac{2u\alpha}{\kappa R} \frac{K_0(x)}{K_1(\kappa R)} \right] - \frac{2u\alpha}{\kappa R} \frac{K_0(x)}{K_1(\kappa R)} \right\} dx \quad (15)$$

The solution  $\alpha(\kappa R)$ , where  $R \simeq 1$  nm is the DNA radius, is plotted in Figure 6 for  $z = 1$  (red) and  $z = 2$  (blue). The



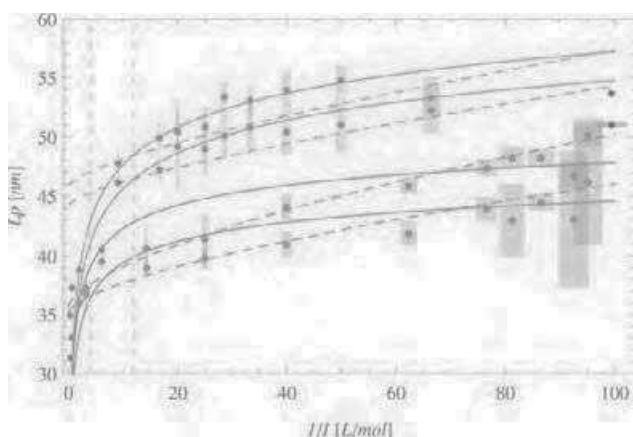
**Figure 6.** Renormalization charge parameter  $\alpha$  vs dimensionless screening parameter  $\kappa R$ , solution of eq 15 for  $u = 4.11$  (monovalent counterions in red) and  $u = 8.23$  (divalent counterions in blue). The black circles correspond to the experimental values studied in section 3.

renormalization factor  $\alpha$  is a monotonous increasing function of  $\kappa R$ , with  $\alpha(\kappa R \rightarrow \infty) \rightarrow 1$  and  $\alpha(\kappa R \rightarrow 0) \rightarrow 1/u$ . Hence it induces a concave shape to  $L_p$ , defined as

$$L_p^{\text{MF}} = L_p^0 + \frac{l_B}{4A^2\kappa^2} [\alpha(\kappa R)]^2 \quad (16)$$

One observes in Figure 6 that the variations of  $\alpha(\kappa R)$  are greater for  $z = 2$  than for  $z = 1$ . Moreover, for  $\kappa R \ll 1$ ,  $\alpha$  for  $z = 2$  is smaller than for  $z = 1$ . These two features are in qualitative agreement with what is observed in Figure 7.

To do a quantitative comparison of eq 16 with the experimental data in the whole range of ionic strengths, we used a polynomial interpolation function to fit  $L_p$  for monovalent ions ( $z = 1$ ), and a power law for divalent ions ( $z = 2$ ),  $\alpha(\kappa R) \simeq 0.423(\kappa R)^{0.364}$ , shown in Figure 6. The fitting parameters are  $L_p^0$  and a prefactor in front of the second term on the right-hand-side of eq 16 (expected to be close to 1). Fits are shown in Figure 7 as dashed lines. Clearly, this approach leads to a slightly concave curve for  $L_p(I^{-1})$  for the two types of counterions. However, whereas fits of  $\text{Mg}^{2+}$  data (in blue) are reasonably good ( $L_p^0 = 35.7$  nm, prefactor equal to 1.8 for  $L = 2060$  bp, and  $L_p^0 = 35.2$  nm, prefactor of 1.5 for  $L = 1201$  bp),



**Figure 7.** Same as Figure 5, where dashed lines are fits using eq 16 and solid lines using eq 19. Parameter values are given in the text.

the fits of the  $\text{Na}^+$  data are not good for  $I > 0.1$  mol/L ( $\Gamma^{-1} < 10$  L/mol), the concavity being not pronounced enough ( $L_p^0 = 44.2$  nm, prefactor of 0.8 for  $L = 2060$  bp, and  $L_p^0 = 45.9$  nm, prefactor of 0.9 for  $L = 1201$  bp). Moreover, here again the non-electrostatic contribution to the persistence length,  $L_p^0$ , varies from the fit of  $\text{Na}^+$  data to the fit of the  $\text{Mg}^{2+}$  data.

**5.4. Beyond Mean-Field: Ion–Ion Correlations and Thermal Fluctuations.** Other approaches have been proposed that go beyond the mean-field approximation by taking into account ion–ion correlations or/and thermal fluctuations.<sup>52–54</sup>

Nguyen et al.<sup>54</sup> considered ion–ion correlation in the strong coupling regime, defined as  $\Gamma = z^{3/2}l_B/(2RA)^{1/2} \gg 1$ . The theoretical limit  $\Gamma \rightarrow \infty$  corresponds to the freezing of the strongly coupled counterions into a Wigner crystal close to the DNA molecule. For  $\Gamma \gg 1$ , they found the following correction to the persistence length in the limit of zero ionic strength:

$$L_p = L_p^0 + \frac{R^2}{2zA}(-0.83\Gamma + 0.33\Gamma^{1/4} + 0.87) \quad (17)$$

which is independent of the ionic strength  $I$ . Therefore, they propose that  $L_p$  is constant at vanishing  $I$ . For DNA at room temperature, one has  $\Gamma = 1.2z^{3/2}$ , and applying eq 17 yields a correction of +0.65 nm for  $z = 1$  (even if  $\Gamma \simeq 1$ ), and –2.20 nm for  $z = 2$ . Equation 17 therefore qualitatively explains the observed abrupt decrease of  $L_p$  by about 5 nm when  $\text{Mg}^{2+}$  ions are added at very low  $I$  in the buffer: the monovalent  $\text{Na}^+$  ions are replaced by strongly coupled divalent counterions, whose correlations decrease the global free energy of the condensed ions and therefore the DNA bending free energy. Hence,  $L_p(I \rightarrow 0)$  decreases with  $z$ . Following this approach, it is thus consistent to choose two different asymptotic values for  $L_p$  when  $I \rightarrow 0$  for monovalent and divalent ions. This explains why these values were slightly different (by about 9 nm) in the preceding section. Note that this theory explains the constant shift at very low  $I$  but does not explain the change in the shape of  $L_p(I^{-1})$ .

Thermal fluctuations were taken into account by Golestanian et al.,<sup>53</sup> who obtained a correction to the persistence length due to fluctuation-induced correlations between ions. Indeed, they correct the OSF–Manning formula, eq 11, at low  $\kappa A$ , following

$$L_p = L_p^0 + \frac{l_B}{4u^2(\kappa A)^2}f(\kappa A, u) \quad (18)$$

where  $f(\kappa A, u) = [1 - 2(u - 1) \ln(\kappa A)]^{-2}$ . Ariel and Andelman<sup>52</sup> proposed a similar correction, with  $f(\kappa A, u) = u(2 - u) - (u - 1)^2/[u \ln(\kappa A)]$ .

Both formulas do not apply to DNA for  $z = 1$  and 2 ( $u = 4.21$  and 8.41). Golestanian's formula yields a slope that is divided by almost 200, with a convex shape for  $\Gamma^{-1} < 10$  L/mol, and Ariel's formula yields a decreasing function of  $L_p(I^{-1})$  as  $\Gamma^{-1}$  increases for the whole  $I$  range. These two theories are therefore not consistent with the whole set of experimental data shown in section 3.

**5.5. Manning's Internal Stretching Force Calculation for  $L_p$ .** The major issue in trying to fit the above theories for the whole range of ionic strengths is to find a fit that yields the concave shape observed for the experimental values. Equation 7, *a priori* valid for high  $I$ , and eq 11, valid for low  $I$ , cannot be reconciled because the constant value is clearly different in both cases,  $L_p^0 > L_p^\infty$ , and should therefore vary with  $I$ .

Manning noted this discrepancy in 2006,<sup>15</sup> and proposed a new formula for the persistence length taking into account the internal electrostatic tension due to the repulsion between charges along the polyelectrolyte. He adapted the calculation reported by Netz<sup>55</sup> for strongly stretched polyelectrolytes at the Debye–Hückel level to the framework of the counterion condensation approach to obtain the persistence length of a polyelectrolyte as a function of  $\kappa A$  and the persistence length of the so-called null isomer (the hypothetical structure of the polyelectrolyte if the backbone charges are set to zero),  $L_p^*$ :

$$L_p = \left(\frac{\pi}{2}L_p^*\right)^{2/3} \frac{R^{4/3}}{z^2l_B} \left[ (2u - 1) \frac{\kappa A e^{-\kappa A}}{1 - e^{-\kappa A}} - 1 - \ln(1 - e^{-\kappa A}) \right] \quad (19)$$

Equation 19 fits very well our data for the  $\text{Na}^+$  case over the whole ionic range with only one fitting parameter,  $L_p^*$  (Figure 7). The fitting values are  $L_p^* = 6.0$  nm (for  $L = 2060$  bp) and 5.4 nm (for  $L = 1201$  bp), close to the value of 7.4 nm fitted by Manning using various experimental data for long DNA ( $L \simeq 40$  kbp).<sup>15</sup> This suggests the phenomenological scaling  $L_p^*(L) \simeq 3.07L^{0.083}$ .

Savelyev<sup>4</sup> performed numerical simulations to investigate the dependence of the persistence length of double-stranded DNA on solutions with various ionic strengths. A coarse-grained model of a two-bead DNA chain with explicit mobile ions ( $\text{Na}^+$  and  $\text{Cl}^-$  ions)<sup>16</sup> was designed to reproduce physical salt conditions from  $10^{-4}$  to 0.1 mol/L (the water solvent is implicit). Their numerical results for  $L_p(I)$  are in semi-quantitative agreement with eq 19 for  $I > 0.1$  mol/L. For lower  $I$ , the agreement is better with the OSF theory, eq 7 (see Figure 2 of ref 4; the fitting parameter values are not given). This was confirmed in other molecular dynamics simulations,<sup>56</sup> where the non-electrostatic contribution to  $L_p$  (with the DNA charges set to 0) appeared to be much larger than the  $L_p^*$  value found by Manning. Moreover, Savelyev<sup>4</sup> compared previous experimental results to his simulations and found a qualitative agreement.

Assuming that, according to the counterion condensation theory,<sup>44</sup> all the condensed counterions are divalent, we use the same formula for the persistence lengths with  $\text{Mg}^{2+}$  counterions. It leads to poorer fits (Figure 7) with very different values for  $L_p^*$ : 14.1 nm (for  $L = 2060$  bp) and 12.7 nm (for  $L = 1201$  bp). The fact that  $L_p^*$  varies, and increases, with the counterion



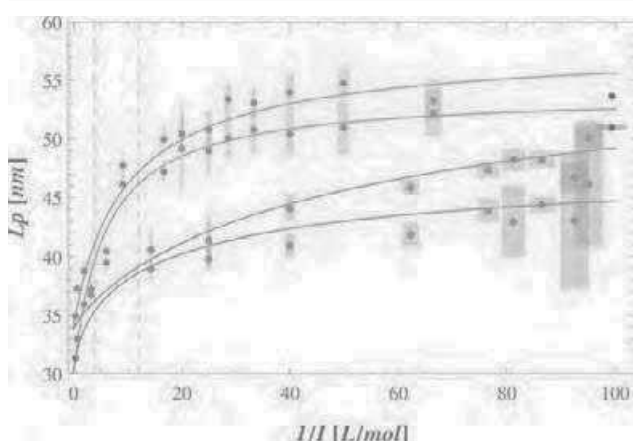
valency is puzzling. Hence, the dependence of  $L_p$  on  $z$  in eq 19 is not consistent with our experimental data.

## 6. INTERPOLATION FORMULA FOR THE WHOLE IONIC STRENGTH RANGE

We propose the following interpolation formula to fit the four sets of data (two DNA lengths, monovalent and divalent salts) over the whole  $I$  range:

$$L_p = L_p^\infty + \frac{L_p^0 - L_p^\infty}{1 + (I/I_0)^\delta} \quad (20)$$

with four fitting parameters:  $L_p^0$ ,  $L_p^\infty$ ,  $I_0$ , and  $\delta$ . The fits, shown in Figure 8, are very good for the monovalent  $\text{Na}^+$  ion, with  $L_p^0 =$



**Figure 8.** DNA persistence length,  $L_p$ , vs the inverse of the ionic strength,  $I^{-1}$ , extracted from the HT-TPM data. The red (resp. blue) symbols correspond to buffers with sodium (resp. magnesium) counterions (top curves are for  $L = 2060$  bp and bottom curves for  $L = 1201$  bp). The solid lines are fits using eq 20 with parameter values given in the text.

58.1 nm,  $L_p^\infty = 33.8$  nm,  $I_0 = 0.104$  mol/L, and  $\delta = 0.931$  for  $L = 2060$  bp, and  $L_p^0 = 53.9$  nm,  $L_p^\infty = 30.1$  nm,  $I_0 = 0.174$  mol/L, and  $\delta = 0.994$  for  $L = 1201$  bp.

Several comments are in order. First, as expected, the asymptotic values,  $L_p^\infty$  and  $L_p^0$ , are, at about 2 nm, the same as those extracted from the linear OSF and OSF–Manning fits shown in Figure 5. Next, the value of the crossover ionic strength,  $I_0$ , is on the order of 0.1 mol/L, which corresponds to a Debye screening length  $\kappa_0^{-1} \simeq 1$  nm, i.e., on the order of the DNA radius,  $R$ . There thus suggest that the concave shape, which is more pronounced for  $I \simeq I_0$ , comes from nonlinear Poisson–Boltzmann effects, as illustrated in section 5.3. Moreover, the effective power law for  $I \simeq I_0$  is found by doing a logarithmic expansion of eq 20 around  $I_0$ :

$$\ln(L_p - L_p^\infty) \simeq \ln\left(\frac{L_p^0 - L_p^\infty}{2}\right) - \frac{\delta}{2} \ln\left(\frac{I}{I_0}\right) \quad (21)$$

which yields an exponent  $-\delta/2 \simeq 0.5$ , which is a good approximation for  $0.05 < I < 0.5$  mol/L (error less than 1 nm). Finally, at low ionic strength,  $I \ll I_0$ , eq 20 varies slightly, and the curve looks like a linear law as a function of  $I^{-1}$  with a small slope, as suggested by eq 11. At large ionic strength,  $I \gg I_0$ , eq 20 yields  $L_p \simeq L_p^\infty + (L_p^0 - L_p^\infty)(I_0/I)^\delta$ , which is equivalent to eq 7 for  $\delta = 1$ , but with a slightly larger slope of 2–4 nm·L/mol.

For divalent  $\text{Mg}^{2+}$  ions, the fits are also good, but with different values for  $\delta$  and  $I_0$  as compared to the  $\text{Na}^+$ -only case. The parameter values  $L_p^\infty$  are equal to those in the  $\text{Na}^+$  case, and the values of  $L_p^0$  are comparable:  $L_p^0 = 59.2$  nm,  $I_0 = 0.017$  mol/L, and  $\delta = 0.830$  for  $L = 2060$  bp, and  $L_p^0 = 51.7$  nm,  $I_0 = 0.045$  mol/L, and  $\delta = 0.546$  for  $L = 1201$  bp. As expected, the concavity is less pronounced and shifted to lower values of  $I$ , close to  $I_0$ . Thus, eq 20 can be useful for experimentalists to interpolate values of  $L_p$  over the whole range of  $I$ .

## 7. CONCLUSION

**7.1. Summary.** Using the high-throughput tethered particle motion setup, we measured the impact of the ionic strength on DNA conformation for two DNA samples of lengths 2060 and 1201 bp. To this end, we investigated a large and homogeneously distributed range of ionic strengths,  $I \in [0.01, 3]$  mol/L, by adding salt to the buffer with monovalent  $\text{Na}^+$  or divalent  $\text{Mg}^{2+}$  counterions. Experimental drift and biases due to the finite exposure time of the detector were corrected. To extract properly the DNA persistence length,  $L_p$ , from the HT-TPM amplitude of motion,  $R_{\text{exp}}$ , numerical exact sampling simulations (without explicit mobile ions and solvent) were performed. Both the DNA excluded volume and that of the particle were taken into account. These simulations allowed us to obtain the experimental  $L_p$  as a function of  $I$  with a good accuracy of about 4% (Figure 7). When  $L_p$  is plotted as a function of  $I^{-1}$ , the overall trend is a monotonous increasing function with, for the  $\text{Na}^+$  case, a concave shape, and, for the  $\text{Mg}^{2+}$  one, an almost linear shape (except at very low concentrations of  $\text{Mg}^{2+}$ ).

Our results are compared to other results found in the literature. A quantitative comparison is difficult, since the  $L_p$  values fluctuate appreciably depending on the experimental setup and the method of extraction of  $L_p$ . Hence, for instance at  $I \simeq 150$  mmol/L,  $L_p$  lies between 40 and 74 nm (Table 1). The available experimental values can, as proposed by Savelyev et al.,<sup>4</sup> be divided into two sets of data. The global behavior of our measured  $L_p$  with  $I$  is not coherent with the first set nor with the second set of experiments. On one hand, our  $L_p$  values appear to be in agreement with the slow increase (of 10%) of  $L_p$  observed by the first group at low  $I$ . On the other hand, at high  $I$ , our  $L_p$  values show a significant 25% variation, in perfect agreement with the experiments of the second group.

Our experimental  $L_p$  values follow a linear OSF law in  $I^{-1}$  only for a very small range of ionic strengths at high  $I$ , with a different prefactor than predicted by OSF in eq 7.

For monovalent  $\text{Na}^+$  counterions,  $L_p$  varies linearly with  $I^{-1}$  at very low  $I$ , according to the OSF equation using the Manning counterion condensation theory, eq 11. The whole  $I$  range is furthermore well fitted by eq 19, which takes into account both the DNA internal stretching due to phosphate ions of the backbone and the counterion condensation around the DNA.

For divalent  $\text{Mg}^{2+}$  counterions, however, neither eq 11 nor eq 19 (with the same fitting parameter as for  $\text{Na}^+$ ) fits well the data. This suggests that these theories do not reproduce well the observed dependence on the valency  $z$ . Using a variational approach taking into account both nonlinear Poisson–Boltzmann effects and screening by mobile ions, we propose a reasonable fit for both the  $\text{Na}^+$  and the  $\text{Mg}^{2+}$  cases, but only for  $I < 0.1$  mol/L and with two different values of  $L_p^0$  at vanishing ionic strength. This marked decrease of  $L_p^0$  at very low  $I$ , when a very small amount of  $\text{Mg}^{2+}$  ions is added, is

semiquantitatively explained by a theory that considers ion–ion correlations, eq 17, for large  $z$ .

In order to both interpolate  $L_p$  values between the ones effectively measured and compare with future experimental data, we proposed an empirical formula that fits both the monovalent and divalent cases.

**7.2. Concluding Remarks.** The large scattering of the available experimental  $L_p$  values observed in Figure 4 may be due to the different experimental setups but also to the various buffers used in these experiments. Indeed, we have shown that the presence of traces of divalent ions, which are often present in buffers in order to maintain the fixed pH, can substantially decrease the  $L_p$  value at a given ionic strength.

An illustration of the extreme sensitivity of persistence length values to the experimental method and the model employed to extract the results was shown by Mielke et al.<sup>57</sup> Brownian dynamics simulations on a double-stranded DNA in a bulk environment were performed at different salt concentrations. Two different strategies were employed to calculate  $L_p$  from the simulation results. One used the expression of the WLC model, eq 1, and the other used an approximation proposed by Hagerman<sup>2</sup> for the rotational diffusion coefficients in order to directly connect to the experimental results of ref 27. At low concentrations, depending on the method used,  $L_p(I)$  shows two distinct behaviors. Values from the rotational diffusion coefficients were more than 30 times larger than the WLC values and Hagerman's values. This result highlights the effects of the chosen model to extract  $L_p$  and the rough approximation used in earlier models to extract it from DLS measurements.

Furthermore, it was recently shown that HT-TPM can detect the effect of the DNA sequence, in particular the presence of A-tracts, on the DNA conformation, which has been interpreted as a modification of the DNA's spontaneous curvature.<sup>19</sup> Preliminary results also show that, for a given DNA length but two different sequences, the persistence length varies. It is thus tempting to suggest that the bare, non-electrostatic contribution to  $L_p$  can also be sequence-dependent, and this can be another explanation for the data scattering.

Finally, many experiments study the influence of ions with higher valency—for instance, trivalent ions such as spermidine, which also has a strong effect on  $L_p$  at millimolar concentrations<sup>7,36</sup>—and/or the role of multivalent ions on the DNA melting temperature.<sup>58</sup> It would be interesting to pursue such a quantitative study of the DNA conformation for such trivalent ions, and thus to study the interplay between screening effects, condensation,<sup>59</sup> and denaturation.<sup>60</sup>

On the theoretical side, a complete theory that explains the variations of  $L_p$  as a function of both  $I$  and the counterion valency  $z$  is still lacking. The approach used by Manning<sup>15</sup> is appealing since it fits very well the experimental data for monovalent counterions with only one fitting parameter. However, the precise treatment of ion–ion correlations should be taken into account to extend such a theory to counterions with higher valencies. Note that these approaches do not consider dielectric exclusion close to a low dielectric molecule such as DNA<sup>61,62</sup> or van der Waals interactions, which are also modified when the ionic strength is varied.

In this paper, we assumed that, according to the WLC model, the tangent–tangent correlation function is a simple exponential, eq 1, which therefore leads to a single correlation length,  $L_p$ . However, Barrat and Joanny show that, by taking into account the polymer bending fluctuations, the persistence length is scale-dependent.<sup>63</sup> A more appropriate choice would

be a double exponential where  $\langle \mathbf{t}(s) \cdot \mathbf{t}(0) \rangle \simeq 1 - s/L_p^\infty$  at small length scales and  $\langle \mathbf{t}(s) \cdot \mathbf{t}(0) \rangle = \exp(-s/L_p^{\text{OSF}})$  at large length scales, where the crossover depends on the ionic strength. It has been shown<sup>64</sup> that  $L_p^{\text{OSF}}$  is given by eq 7 and  $L_p^\infty$  is the bare persistence length. Another approach proposed that the latter is also salt-dependent.<sup>65</sup> Of course, such a model is more difficult to apply to the experimental measure of  $R_{\text{DNA}}$  only, with a rather tricky extraction of two different correlation lengths, but it might provide a relevant framework to explain the overall observed behaviors of  $L_p(I)$ .

## ■ ASSOCIATED CONTENT

### Supporting Information

Plots of DNA approximate end-to-end distances, tables of experimental data, and theoretical complements. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.macromol.5b00735.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: manghi@irsamc.ups-tlse.fr.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank Oriol Servera Sires for his valuable B.Sc. work on the fitting procedure, and Juliette Wilhem who helped with the experiments as part of her Master's project. We acknowledge financial support from the CNRS, University of Toulouse 3, and ANR-11-NANO-010 TPM-On-a-Chip.

## ■ REFERENCES

- (1) Harrington, R. E. *Biopolymers* **1978**, *17*, 919.
- (2) Hagerman, P. J. *Biopolymers* **1981**, *20*, 1503.
- (3) Borochov, N.; Eisenberg, H.; Kam, Z. *Biopolymers* **1981**, *20*, 231.
- (4) Savelyev, A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 2250.
- (5) Rizzo, V.; Schellmann, J. *Biopolymers* **1981**, *20*, 2143.
- (6) Maret, G.; Weill, G. *Biopolymers* **1983**, *22*, 2727.
- (7) Baumann, C. G.; Smith, S. B.; Bloomfield, V. A.; Bustamante, C. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 6185.
- (8) Manning, G. S. *Biopolymers* **1981**, *20*, 1751.
- (9) Kam, Z.; Borochov, N.; Eisenberg, H. *Biopolymers* **1981**, *20*, 2671.
- (10) Cairney, K. L.; Harrington, R. E. *Biopolymers* **1982**, *21*, 923.
- (11) Post, C. B. *Biopolymers* **1983**, *22*, 1087.
- (12) Sobel, E. S.; Harpst, J. A. *Biopolymers* **1991**, *31*, 1559.
- (13) Odijk, T. J. *Polym. Sci.* **1977**, *15*, 477.
- (14) Skolnick, J.; Fixman, M. *Macromolecules* **1977**, *10*, 944.
- (15) Manning, G. S. *Biophys. J.* **2006**, *91*, 3607.
- (16) Savelyev, A.; Papoian, G. A. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 171.
- (17) Peters, J. P.; Maher, L. J., III, Q. *Rev. Biophys.* **2010**, *43*, 23.
- (18) Plenat, T.; Tardin, C.; Rousseau, P.; Salomé, L. *Nucleic Acids Res.* **2012**, *40*, e89.
- (19) Brunet, A.; Chevalier, S.; Destainville, N.; Manghi, M.; Rousseau, P.; Salhi, M.; Salomé, L.; Tardin, C. *Nucleic Acids Res.* **2015**, DOI: 10.1093/nar/gkv201.
- (20) Pouget, N.; Dennis, C.; Turlan, C.; Grigoriev, M.; Chandler, M.; Salomé, L. *Nucleic Acids Res.* **2004**, *32*, e73.
- (21) Manghi, M.; Tardin, C.; Baglio, J.; Rousseau, P.; Salomé, L.; Destainville, N. *Phys. Biol.* **2010**, *7*, 046003.
- (22) Destainville, N.; Salomé, L. *Biophys. J.* **2006**, *90*, L17.
- (23) Segall, D. E.; Nelson, P.; Phillips, R. *Phys. Rev. Lett.* **2006**, *96*, 4.
- (24) Doi, M.; Edwards, S. F. *The Theory of Polymer Dynamics*; Oxford University Press: Oxford, 1986.

- (25) Yamakawa, H. *Helical Wormlike Chains in Polymer Solutions*; Springer-Verlag: Berlin, 1997.
- (26) Palmeri, J.; Manghi, M.; Destainville, N. *Phys. Rev. E* **2008**, *77*, 011913.
- (27) Hagerman, P. J. *Ann. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 265.
- (28) Porschke, D. *Biophys. Chem.* **1991**, *40*, 169.
- (29) Elias, J. G.; Eden, D. *Macromolecules* **1981**, *14*, 410.
- (30) Elias, J. G.; Eden, D. *Biopolymers* **1981**, *20*, 2369.
- (31) Diekmann, S.; Hillen, W.; Morgeneyer, B.; Wells, R. D.; Porschke, D. *Biophys. Chem.* **1982**, *15*, 263.
- (32) Porschke, D. *J. Biomol. Struct. Dyn.* **1986**, *4*, 373.
- (33) Bednar, J.; Furrer, P.; Katritch, V.; Stasiak, A.; Dubochet, J.; Stasiak, A. *J. Mol. Biol.* **1995**, *254*, 579.
- (34) Lu, Y.; Weers, B.; Stellwagen, N. C. *Biopolymers* **2002**, *61*, 261.
- (35) Borochoy, N.; Eisenberg, H. *Biopolymers* **1983**, *23*, 1757.
- (36) Wang, M. D.; Yin, H.; Landick, R.; Gelles, J.; Block, S. M. *Biophys. J.* **1997**, *72*, 1335.
- (37) Mantelli, S.; Muller, P.; Harlepp, S.; Maaloum, M. *Soft Matter* **2011**, *7*, 3412.
- (38) Smith, S. B.; Finzi, L.; Bustamante, C. *Science* **1992**, *258*, 1122.
- (39) Dietrich, H. R. C.; Rieger, B.; Wiertz, F. G. M.; De Groote, F. H.; Heering, H. A.; Young, I. T.; Garini, Y. *J. Nanophoton.* **2009**, *3*, 031795.
- (40) Han, L.; Lui, B. H.; Phillips, R. In *Mathematics of DNA Structure, Function and Interactions*; Benham, C., Harvey, J. S., Olson, W. K., Summers, D. W. L., Swigon, D., Eds.; Springer: New York, 2009; pp 123–138.
- (41) Kumar, S.; Manzo, C.; Zurla, C.; Ucuncuoglu, S.; Finzi, L.; Dunlap, D. *Biophys. J.* **2014**, *106*, 399.
- (42) Fixman, M. *J. Phys. Chem. B* **2010**, *114*, 3185.
- (43) Manning, G. S. *J. Chem. Phys.* **1969**, *51*, 924.
- (44) Manning, G. S. *J. Chem. Phys.* **1984**, *88*, 6654.
- (45) Le Bret, M. *J. Phys. Chem.* **1982**, *76*, 6243.
- (46) Fixman, M. *J. Phys. Chem.* **1982**, *76*, 6346.
- (47) Wenner, J. R.; Williams, M. C.; Rouzina, I.; Bloomfield, V. A. *Biophys. J.* **2002**, *82*, 3160.
- (48) Tomić, S.; Vuletić, T.; Dolanski Babić, S.; Krča, S.; Ivanković, D.; Griparić, L.; Podgornik, R. *Phys. Rev. Lett.* **2006**, *97*, 098303.
- (49) Grgičin, D.; Dolanski Babić, S.; Ivek, T.; Tomić, S.; Podgornik, R. *Phys. Rev. E* **2013**, *88*, 052703.
- (50) Manning, G. S. *Macromolecules* **2001**, *34*, 4650.
- (51) Netz, R. R.; Orland, H. *Eur. Phys. J. E* **2003**, *11*, 301.
- (52) Ariel, G.; Andelman, D. *EPL* **2003**, *61*, 67.
- (53) Golestanian, R.; Kardar, M.; Liverpool, T. *Phys. Rev. Lett.* **1999**, *82*, 4456.
- (54) Nguyen, T. T.; Rouzina, I.; Shklovskii, B. *Phys. Rev. E* **1999**, *60*, 7032.
- (55) Netz, R. R. *Macromolecules* **2001**, *34*, 7522.
- (56) Savelyev, A.; Materese, C. K.; Papoian, G. A. *J. Am. Chem. Soc.* **2011**, *133*, 19290.
- (57) Mielke, S. P.; Benham, C. J.; Grønbech-Jensen, N. *J. Phys. Chem. A* **2009**, *113*, 4213.
- (58) Owczarzy, R.; Moreira, B. G.; You, Y.; Behlke, M. A.; Walder, J. A. *Biochemistry* **2008**, *47*, 5336.
- (59) Bloomfield, V. A. *Curr. Opin. Struct. Biol.* **1996**, *6*, 334.
- (60) Korolev, N.; Lyubartsev, A. P.; Nordenskiöld, L. *Biophys. J.* **1998**, *75*, 3041.
- (61) Chertsvy, A. G. *J. Phys. Chem. B* **2007**, *111*, 12933.
- (62) Buyukdagli, S.; Manghi, M.; Palmeri, J. *Phys. Rev. E* **2010**, *81*, 041601.
- (63) Barrat, J.-L.; Joanny, J.-F. *Europhys. Lett.* **1993**, *24*, 333.
- (64) Manghi, M.; Netz, R. R. *Eur. Phys. J. E* **2004**, *14*, 67.
- (65) Gubarev, A.; Carrillo, J.-M. Y.; Dobrynin, A. V. *Macromolecules* **2009**, *42*, 5851.
- (66) Ramanathan, G. V. *J. Chem. Phys.* **1983**, *78*, 3223.
- (67) Stellwagen, N. C. *Biopolymers* **1981**, *20*, 399.

# Supplemental Information

## Dependence of DNA persistence length on ionic strength of solutions with monovalent and divalent salts: a joint theory-experiment study

Annaël Brunet,<sup>1,2,3,4</sup> Catherine Tardin,<sup>1,2</sup> Laurence Salomé,<sup>1,2</sup>  
Philippe Rousseau,<sup>5,6</sup> Nicolas Destainville,<sup>3,4</sup> and Manoel Manghi<sup>3,4</sup>

<sup>1</sup>*Institut de Pharmacologie et Biologie Structurale,  
UMR CNRS-UPS 5089, 31062 Toulouse, France*

<sup>2</sup>*Université de Toulouse, UPS, Institut de Pharmacologie et Biologie Structurale, 31062 Toulouse, France*

<sup>3</sup>*Université de Toulouse, UPS, Laboratoire de Physique Théorique (IRSAMC), 31062 Toulouse, France*

<sup>4</sup>*CNRS, Laboratoire de Physique Théorique (IRSAMC), 31062 Toulouse, France*

<sup>5</sup>*Université de Toulouse, UPS, LMGM (Laboratoire de Microbiologie et Génétique Moléculaires), 31062 Toulouse, France*

<sup>6</sup>*CNRS, LMGM, UMR CNRS-UPS 5100, 31062 Toulouse, France*

(Dated: April 9, 2015)

### I. DNA END-TO-END DISTANCE

The DNA end-to-end distance,  $R_{\text{DNA}}$  extracted from the raw HT-TPM amplitudes of motion of Fig. 1 of the article using Eq. (3) are shown in Fig. 1.

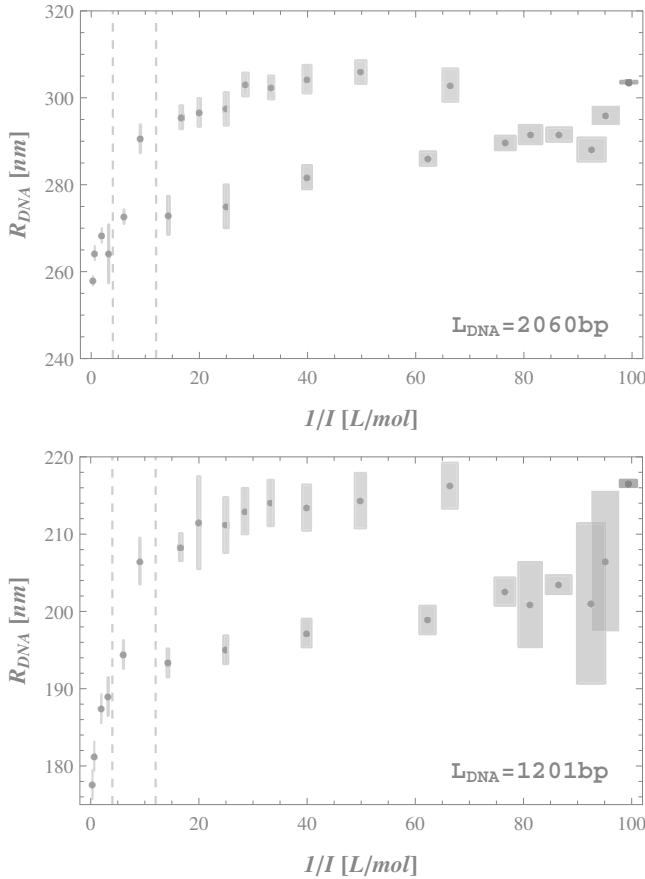


FIG. 1: DNA approximate end-to-end distance [extracted using Eq. (3) of the article] vs.  $I^{-1}$  for a DNA of length (a)  $L = 2060$  bp DNA and (b)  $L = 1201$  bp with both  $\text{Na}^+$  (red) and  $\text{Mg}^{2+}$  (blue) cations.

### II. VARIATIONAL CALCULATION OF $\alpha(\kappa)$

The dimensionless electrostatic Debye-Hückel potential,  $\phi = ze\beta\psi$ , for an electrolyte (valency  $z$ ) of Debye-Hückel constant  $\kappa$  around a cylinder of radius  $R$  and surface charge density  $\sigma = (2\pi AR)^{-1}$  empty of ions is:

$$\phi(r) = \frac{2u}{\kappa R} \frac{K_0(\kappa r)}{K_1(\kappa R)} \quad \text{for } r > R \quad (1)$$

$$\phi(r) = \frac{2u}{\kappa R} \frac{K_0(\kappa R)}{K_1(\kappa R)} \quad \text{for } r \leq R \quad (2)$$

where  $K_0$  and  $K_1$  are the modified Bessel function of order 0 and 1 and  $u = z\ell_B/A$  is the Manning parameter.

Following Netz and Orland [1], the full non-linear Poisson-Boltzmann equation is solved variationally by assuming that the solution is of the Debye-Hückel form,  $\alpha\phi(r)$ , where  $\phi(r)$  is given in Eq. (1), and  $\alpha$ , the fraction of the “free” counter-ions, is the variational parameter, and is given by:

$$4\pi z\ell_B(1-\alpha) \int d\mathbf{r} \rho(\mathbf{r}) \phi(\mathbf{r}) = \kappa^2 \int d\mathbf{r} \Omega(\mathbf{r}) \phi(\mathbf{r}) \{ \sinh[\alpha\phi(\mathbf{r})] - \alpha\phi(\mathbf{r}) \} \quad (3)$$

where  $\rho(\mathbf{r}) = \sigma\delta(r-R)$  is the charge distribution. Note that the ionic exclusion factor  $\Omega(\mathbf{r})$  was incorrectly put just in front of the sinh in [1]. Using Eq. (1), Eq. (3) simplifies to

$$2u(1-\alpha)K_0(\kappa R) = \int_{\kappa R}^{\infty} x K_0(x) \left\{ \sinh \left[ \frac{2u\alpha}{\kappa R} \frac{K_0(x)}{K_1(\kappa R)} \right] - \frac{2u\alpha}{\kappa R} \frac{K_0(x)}{K_1(\kappa R)} \right\} dx \quad (4)$$

An asymptotic analysis of Eq. (4) leads to:

$$\alpha \simeq 1 - \frac{u^2}{6(\kappa R)^2} \quad \text{for } \kappa R \gg 1 \quad (5)$$

In particular, as discussed in [1], we reach  $\alpha = 1$  at high ionic strengths. Note that the renormalization charge vanishes with a different prefactor,  $1/6$  instead of  $1/24$ .

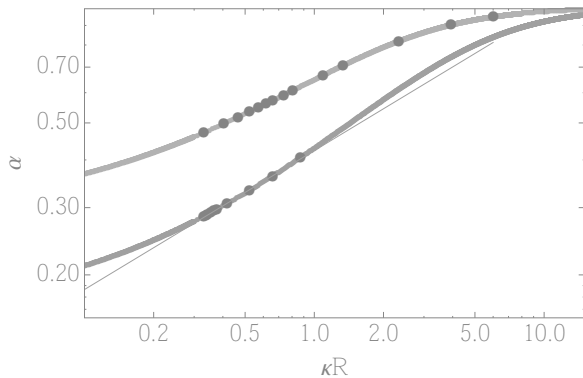


FIG. 2: Renormalization charge parameter  $\alpha$  vs. dimensionless screening parameter  $\kappa R$ , solution of Eq. (4) for  $u = 4.11$  (monovalent counter-ions in red) and  $u = 8.23$  (divalent counter-ions in blue). The black circles correspond to the experimental values studied in the article.

Indeed, since we assume that the mobile ions to be present only on one side of the charged surface, according to the Gauss law the electrostatic field (and potential) is twice the one for the case of a surface penetrable for ions. Hence for the planar case, the potential is [2],  $\phi(z) = 2e^{-\kappa z}/\tilde{\kappa}$ , where  $\tilde{\kappa} = \kappa/(2\pi z\ell_B\sigma)$ , and Eq. (38) of Ref. [1] is replaced by

$$1 - \frac{\alpha}{2} = \frac{\tilde{\kappa}^2}{4\alpha} \left[ \cosh\left(\frac{2\alpha}{\tilde{\kappa}}\right) - 1 \right] \quad (6)$$

which leads to  $\alpha = 1 - 1/(6\tilde{\kappa}^2)$  for  $\tilde{\kappa} \gg 1$ , which is the same result as Eq. (5).

For low ionic strengths, a careful asymptotic analysis leads to

$$\alpha \simeq \frac{1}{u} \left[ 1 + \frac{1}{2(\ln(2/\kappa R) - \gamma)} \left( 1 + \frac{4(u-1)}{e(\ln(2/\kappa R) - \gamma)} \right) \right] \quad (7)$$

for  $\kappa R \ll 1$ , which using Eq. (1) leads to

$$\alpha\phi(R) = 2\ln(1/\kappa R) + 1 + 2(\ln 2 - \gamma) + o(1) \quad (8)$$

which, at first order, compares to the PB result [3]:

$$\phi_{\text{PB}}(R) = 2\ln(1/\kappa R) + 2\ln[2(u-1)] + o(1) \quad (9)$$

In Fig. 2 are plotted the numerical solutions,  $\alpha(\kappa R)$ , of Eq. (4) for monovalent ( $u = 4.12$ ) and divalent counter-ions ( $u = 8.23$ , i.e. we assume here that all the condensed counter-ions are divalent). We also add the values obtained for the experimental ionic strengths as large dots.

### III. EXPERIMENTAL DATA

In tables I and II are listed the experimental data for the DNA lengths,  $L = 2060$  bp and  $L = 1201$  bp.

Calculation of the errors on  $I$  values is based on the estimation of the error on the added volume of each component during the buffer preparation step.

In particular, the relaxation time,  $\tau_{||}$ , of the system DNA/particle is extracted from HT-TPM measurements. The error on  $\tau_{||}$  of an ensemble of particles was calculated by using the bootstrap method of the R software, as for  $R_{\text{exp}||}$  (R Foundation for Statistical Computing, Vienna, Austria). Doing so, we find a typical error of 0.4 ms.

We discuss in this appendix the variation of  $\tau_{||}$  with the ionic strength,  $I$ . By adding  $\text{Na}^+$  ions,  $\tau_{||}$  slowly decreases from 52 ms (resp. 36 ms) at  $I = 10$  mmol/L to 49 ms (33 ms) at 510 mmol/L for  $L = 2060$  bp (resp. 1201 bp), even though we cannot exclude this tendency to be due to experimental uncertainties. Such a moderate reduction of the relaxation time was previously observed in TEB studies [4–7]. The decay of the transmitted intensity, proportional to the birefringence or dichroism signals, were directly related to the rotational diffusion constant of the DNA molecules in solution. On one hand, the relationship between light intensity during the field pulse and the optical retardation depends on the optical arrangement. Indeed, the extracted rotational relaxation time is a function of the sample orientation with respect to the polarizer. On the other hand,  $\tau_{||}$  was also independent of the DNA concentration and of electric field strength. Therefore, a standard arrangement was adopted for each study and for all measurements.

In their first studies, Elias *et al.* [8] found a  $\tau_{||}$  independent of  $I$  over a range of 0.2 to 2 mmol/L of  $\text{Na}^+$  concentration. In comparison, on the same short range of  $I$ , we obtain a slow decrease of  $\tau_{||}$ . In a following work [5], the prospected range of  $I$  was extended, and a decrease of  $\tau_{||}$  was observed. Porschke [9] also observed a decrease of the dichroism decay time constant by 20% for a fragment of 84 bp on the range of 2 mmol/L to 100 mmol/L. We did not prospected so small  $I$  values, but on  $I \simeq 10$  mmol/L to 100 mmol/L, we observe a decrease of  $\tau_{||}$  of 6%.

For the last value of  $I = 3$  mol/L, we measured an marked increase of  $\tau_{||}$  from 49 ms (resp. 33 ms) to 57 ms (40 ms) for  $L = 2060$  bp (resp. 1201 bp), which suggests an additional effect on the DNA conformation.

With  $\text{Mg}^{2+}$  ions, the same moderate reduction of the  $\tau_{||}$  was observed. In addition the absolute value of  $\tau_{||}$  was always smaller for divalent than monovalent ions, as for  $L_p$  values. This observation on the absolute values of  $\tau_{||}$  underlines the different impact of monovalent and divalent ions on the DNA dynamics.

- 
- [1] R.R. Netz, H. Orland, Eur. Phys. J. E **11**, 301 (2003)
  - [2] S. Buyukdagli, M. Manghi, J. Palmeri, Phys. Rev. E **81**, 041601 (2010)
  - [3] G.V. Ramanathan, J. Chem. Phys. **78**, 3223 (1983)
  - [4] P.J. Hagerman, Biopolymers **20** 1503 (1981)
  - [5] J.G. Elias, D. Eden, Biopolymers **20**, 2369 (1981)
  - [6] N. C. Stellwagen, Biopolymers **20**, 399-434 (1981)
  - [7] Y. Lu, B. Weers, N.C. Stellwagen, Biopolymers **61**, 261 (2002)
  - [8] J.G. Elias, D. Eden, Macromolecules **14**, 410 (1981)
  - [9] D. Porschke, Biophys. Chem. **40**, 169 (1991)

Ion	$C_{\text{ion}}$ [mmol/L]	$I$ [mmol/L]	$\Delta I$ [mmol/L]	$1/I$ [L/mol]	Nb. traj.	$\tau_{  }$ [ms]	$\Delta\tau_{  }$ [ms]	$R_{\text{exp}  }$ [nm]	$\Delta R_{\text{exp}  }$ [nm]	$R_{\text{DNA}}$ [nm]	$\Delta R_{\text{DNA}}$ [nm]	$L_p$ [nm]	$\Delta L_p$ [nm]
	0	10.1	0.3	99.4	897	51.6	0.3	276.5	0.2	303.6	0.2	53.7	0.2
$\text{Na}^+$	5	15.1	0.3	66.4	204	52.8	0.6	276.0	0.7	302.9	1.9	53.4	0.9
	10	20.1	0.4	49.9	284	52.0	0.5	278.2	0.5	305.9	1.4	54.8	0.7
	15	25.1	0.5	39.9	197	51.1	0.6	277.0	0.6	304.3	1.7	54.0	0.8
	20	30.1	0.5	33.3	230	52.2	0.6	275.6	0.5	302.4	1.7	53.1	0.7
	25	35.1	0.7	28.5	312	51.2	0.6	276.0	0.5	303.1	1.4	53.4	0.7
	30	40.1	0.7	25.0	193	52.0	0.7	272.0	0.7	297.4	1.9	50.8	0.9
	40	50.1	0.8	20.0	223	51.8	0.5	271.4	0.6	296.6	1.7	50.5	0.8
	50	60.1	1.2	16.7	299	51.2	0.2	270.6	0.5	295.5	1.4	50.0	0.6
	100	110.1	1.5	9.1	254	50.2	0.4	267.0	0.6	290.6	1.7	47.8	0.7
	155	165.1	2.7	6.1	1100	49.9	0.2	254.1	0.3	272.7	0.9	40.5	0.4
	500	510.1	7.7	2.0	732	48.6	0.3	251.0	0.3	268.3	0.9	38.8	0.3
	1500	1510.1	20.7	0.7	840	50.9	0.2	248.1	0.3	264.3	0.9	37.3	0.3
	3000	3010.1	37.5	0.3	815	57.1	0.3	243.6	0.2	257.9	0.6	35.0	0.5
$\text{Mg}^{2+}$	0.15	10.5	0.3	95.1	364	49.8	0.3	270.9	0.4	295.9	1.1	50.2	0.5
	0.25	10.8	0.3	92.5	197	48.9	0.4	265.2	0.5	288.1	1.4	46.8	0.6
	0.50	11.6	0.3	86.5	482	48.0	0.4	267.7	0.3	291.5	0.9	48.2	0.4
	0.75	12.3	0.3	81.2	349	47.4	0.3	267.7	0.4	291.5	1.1	48.2	0.5
	1	13.1	0.3	76.6	559	47.9	0.3	266.3	0.3	189.6	0.9	47.4	0.6
	2	16.1	0.4	62.3	524	47.2	0.4	263.7	0.3	286.0	0.9	45.9	0.4
	5	25.1	0.6	39.9	188	46.0	0.7	260.6	0.5	181.7	1.4	44.1	0.4
	10	40.1	0.7	25.0	184	46.4	0.7	255.8	0.9	275.0	2.5	41.4	1.0
	20	70.1	1.2	14.3	156	46.0	0.6	254.3	0.8	273.0	2.3	40.6	0.9
	100	310.1	5.1	3.2	118	45.5	0.6	248.0	1.2	264.1	3.4	37.2	1.3

TABLE I: Salt conditions (nature of ion, ion concentration, ionic strength with error bars), number of HT-TPM trajectories, experimental correlation time  $\tau_{||}$  (with error bars), and experimental results for the particle amplitude of motion  $R_{\text{exp}}$  (with error bars) for DNA length of  $L = 2060$  bp. The corresponding DNA end-to-end distance  $R_{\text{DNA}}$  is computed using Eq. (3) of the article and the persistence length  $L_p$  following the procedure described in the analysis Section (with error bars).

Ion	$C_{\text{ion}}$ [mmol/L]	$I$ [mmol/L]	$\Delta I$ [mmol/L]	$1/I$ [L/mol]	Nb. traj.	$\tau_{  }$ [ms]	$\Delta\tau_{  }$ [ms]	$R_{\text{exp}  }$ [nm]	$\Delta R_{\text{exp}  }$ [nm]	$R_{\text{DNA}}$ [nm]	$\Delta R_{\text{DNA}}$ [nm]	$L_p$ [nm]	$\Delta L_p$ [nm]
	0	10.1	0.3	99.4	1399	35.8	0.2	213.6	0.2	216.6	0.3	51.0	0.1
$\text{Na}^+$	5	15.1	0.3	66.4	215	35.9	0.7	214.9	0.5	216.3	1.5	52.3	1.0
	10	20.1	0.4	49.9	250	35.1	0.5	213.6	0.6	214.3	1.8	51.0	1.2
	15	25.1	0.5	39.9	232	35.4	0.4	213.0	0.5	213.4	1.5	50.4	1.0
	20	30.1	0.5	33.3	231	34.5	0.5	213.4	0.5	214.0	1.5	50.8	1.0
	25	35.1	0.7	28.5	269	34.3	0.5	212.7	0.5	213.0	1.5	50.1	1.0
	30	40.1	0.7	25.0	216	35.2	0.7	211.5	0.6	211.2	1.8	49.0	1.1
	40	50.1	0.8	20.0	55	35.4	0.7	211.7	1.0	211.5	3.0	49.2	1.9
	50	60.1	1.2	16.7	270	33.5	0.2	209.6	0.3	208.3	0.9	47.3	0.6
	100	110.1	1.5	9.1	273	33.0	0.3	208.4	0.5	206.5	1.5	46.2	0.9
	155	165.1	2.7	6.1	869	32.7	0.2	200.5	0.3	194.4	1.0	39.5	0.5
	500	510.1	7.7	2.0	587	32.5	0.2	196.0	0.3	187.4	1.0	36.0	0.5
	1500	1510.1	20.7	0.7	569	35.4	0.2	192.1	0.3	181.3	1.0	33.0	0.5
	3000	3010.1	37.5	0.3	587	39.9	0.2	189.9	0.3	177.6	01.0	31.4	0.5
$\text{Mg}^{2+}$	0.15	10.5	0.3	95.1	42	33.2	1.3	208.4	1.5	206.5	4.6	46.2	2.7
	0.25	10.8	0.3	92.0	33.6	34	0.8	204.8	1.7	201.0	5.2	43.1	2.9
	0.5	11.6	0.3	86.5	714	33.5	0.2	206.4	0.2	203.5	0.6	44.4	0.4
	0.75	12.3	0.3	81.2	83	32.1	0.4	204.7	0.9	200.9	2.8	43.0	1.6
	1	13.1	0.3	76.6	685	32.8	0.2	205.8	0.3	202.6	0.9	43.9	0.5
	2	16.1	0.4	62.3	766	32.1	0.2	203.4	0.3	198.9	0.9	41.9	0.5
	5	25.1	0.6	39.9	636	31.8	0.2	202.3	0.3	197.2	0.9	41.0	0.5
	10	40.1	0.7	25.0	499	31.1	0.3	200.9	0.3	195.0	1.0	39.8	0.5
	20	70.1	1.2	14.3	443	30.7	0.3	299.8	0.3	193.8	1.0	38.9	0.5
	100	310.1	5.1	3.2	348	31.5	0.3	197.0	0.4	189.1	1.3	36.7	0.6

TABLE II: Same as Table I for DNAs of length  $L = 1201$  bp.



## BIBLIOGRAPHIE

- [ALS] D G Alexeev, a a Lipanov, and Skuratovskii IYa. Poly(dA).poly(dT) is a B-type double helix with a distinctively narrow minor groove. *Nature*, 325(6107) :821–823, jan.
- [Fix82] Marshall Fixman. The flexibility of polyelectrolyte molecules. *Journal of chemical physics*, 76(12) :6346–6353, jun 1982.
- [FS78] M. Fixman and J. Skolnick. Polyelectrolyte Excluded Volume Paradox. *Macromolecules*, 11(5) :863–867, sep 1978.
- [HA06] a a Hussain and M E E Abashar. Effect of Ion Sizes on Separation Characteristics of Nanofiltration Membrane Systems. *Most*, 19(1) :1–19, jan 2006.
- [Har78] Rodney E. Harrington. Opticohydrodynamic properties of high-molecular-weight DNA. III. The effects of NaCl concentration. *Biopolymers*, 17(4) :919–936, 1978.
- [HLA83] M Hogan, J LeGrange, and B Austin. Dependence of DNA helix flexibility on base composition. *Nature*, 304(5928) :752–754, aug 1983.
- [Man69a] Gerald S. Manning. Limiting Laws and Counterion Condensation in Polyelectrolyte Solutions. III. An Analysis Based on the Mayer Ionic Solution Theory. *The Journal of Chemical Physics*, 51(8) :3249, sep 1969.
- [Man69b] Gerald S. Manning. Limiting Laws and Counterion Condensation in Polyelectrolyte Solutions. III. An Analysis Based on the Mayer Ionic Solution Theory. *The Journal of Chemical Physics*, 51(8) :3249, sep 1969.
- [Man81] Gerald S. Manning. Limiting Laws and Counterion Condensation in Polyelectrolyte Solutions , 7 . Electrophoretic Mobility and Conductance '. *J Phys Chem*, 185(4) :1506–1515, sep 1981.
- [Man06] Gerald S Manning. The persistence length of DNA is reached from the persistence length of its null isomer through an internal electrostatic stretching force. *Biophysical journal*, 91(10) :3607–3616, nov 2006.
- [Net01] R. R. Netz. Strongly stretched semiflexible extensible polyelectrolytes and DNA. *Macromolecules*, 34(21) :7522–7529, oct 2001.
- [NO03] H. H. Netz and H. Orland. Variational charge renormalization in charged systems. *European Physical Journal E*, 11(3) :301–311, jul 2003.
- [Odi77] T. Odijk. Polyelectrolytes near the rod limit. *Journal of Polymer Science : Polymer Physics Edition*, 15(3) :477–483, mar 1977.
- [OM75] W Olivares and D a McQuarrie. On the theory of ionic solutions. *Biophysical journal*, 15(2 Pt 1) :143–162, feb 1975.
- [Pod04] Rudi Podgornik. Electrostatic contribution to the persistence length of a semiflexible dipolar chain. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 70(3 1) :6243, jun 2004.
- [PW81] L J Peck and J C Wang. Sequence dependence of the helical repeat of DNA in solution. *Nature*, 292(5821) :375–378, jul 1981.

- [Sav12] Alexey Savelyev. Do monovalent mobile ions affect DNA's flexibility at high salt content? *Physical Chemistry Chemical Physics*, 14(7) :2250, 2012.
- [Sch65] C Schildkraut. Dependence of the melting temperature of DNA on salt concentration. *Biopolymers*, 3(2) :195–208, jan 1965.
- [SGP05] Francois Strauss, Claire Gaillard, and Ariel Prunell. Helical Periodicity of DNA, Poly(dA)·Poly(dT) and Poly(dA-dT)·Poly(dA-dT) in Solution. *European Journal of Biochemistry*, 118(2) :215–222, mar 2005.
- [SNP06] Darren E. Segall, Philip C. Nelson, and Rob Phillips. Volume-exclusion effects in tethered-particle experiments : Bead size matters. *Physical Review Letters*, 96(8) :088306, mar 2006.
- [SR90] T Schlapp and G Rödel. Transcription of two divergently transcribed yeast genes initiates at a common oligo(dA-dT) tract. *Molecular & general genetics : MGG*, 223(3) :438–442, sep 1990.
- [UBTC86] L Ulanovsky, M Bodner, E N Trifonov, and M Choder. Curved DNA : design, synthesis, and circularization. *Proceedings of the National Academy of Sciences of the United States of America*, 83(4) :862–866, feb 1986.
- [VCDC96] L Verdone, G Camilloni, E Di Mauro, and M Caserta. Chromatin remodeling during *Saccharomyces cerevisiae* ADH2 gene activation. *Molecular and cellular biology*, 16(5) :1978–1988, may 1996.

## Quatrième partie

# Effet Local : courbure et bulle de dénaturation



# COURBURE LOCALE INTRINSÈQUE À L'ADN

# 6

## SOMMAIRE

INTRODUCTION . . . . .	209
6.1 SÉQUENCES A-TRACT, INDUISANT UNE COURBURE INTRINSÈQUE DANS LA MOLÉCULE D'ADN . . . . .	211
6.1.1 Stratégie de construction des molécules courbées . . . . .	211
6.1.2 Détail des séquences utilisées durant cette étude . . . . .	212
6.2 THÉORIE POUR UNE COURBURE LOCALE . . . . .	212
6.2.1 Model <i>Kinked Worm-Like-Chain</i> . . . . .	213
6.2.2 Paramètres K-WLC pour une courbure au centre de la molécule . . . . .	214
6.2.3 Paramètres K-WLC pour une courbure excentrée . . . . .	214
6.3 MESURE DE LA VARIATION DE LA DISTANCE BOUT-À-BOUT DES COMPLEXES ADN/PARTICULE EN FONCTION DE LA SÉQUENCE . . . . .	215
6.4 EXTRACTION DE LA DISTANCE BOUT-À-BOUT DE L'ADN SEUL . . . . .	215
6.4.1 Simulation MC-D d'un ADN présentant une courbure locale centrée . . . . .	216
6.4.2 Choix du mode d'extraction . . . . .	216
6.5 QUANTIFICATION DE L'ANGLE DE COURBURE À PARTIR D'UN JEU DE CONSTRUCTIONS D'ADN . . . . .	217
6.5.1 Extraction de l'angle de courbure sur un ADN supposé régulier . . . . .	217
6.5.2 Extraction de l'angle de courbure sur un ADN supposé non-régulier . . . . .	218
6.5.3 Prise en compte de l'incertitude due au rayon de la particule . . . . .	219
6.5.4 Comparaison avec les données de la littérature . . . . .	219
6.6 PROPOSITION D'UNE FORMULE SIMPLE AFIN D'ÉVALUER L'ANGLE DE COURBURE À PARTIR DE DEUX MESURES . . . . .	220
6.6.1 Extraction de l'angle de courbure de données de la littérature . . . . .	220
6.6.2 Extraction de l'angle des courbures de la séquence TA <sub>3</sub> T . . . . .	220
Détails des séquences utilisées durant cette étude . . . . .	220
Extraction de la distance bout-a-bout de l'ADN seul . . . . .	221
Estimation de la courbure induite par un TA <sub>3</sub> T . . . . .	221
6.6.3 Extraction de l'angle de courbure de séquences A-tract variées . . . . .	221
Détails des séquences utilisées durant cette étude . . . . .	221
Extraction de la distance bout-à-bout de l'ADN seul . . . . .	222
Estimation de la courbure induite par ces séquences . . . . .	222
6.7 THÉORIE POUR UN INSERT RIGIDE LOCAL . . . . .	222
6.7.1 Modèle <i>Stiffer Worm-Like-Chain</i> . . . . .	222
6.7.2 Argument justifiant cette approche . . . . .	223

6.7.3	Estimation de la variation de la rigidité sur la séquence de A-tracts en opposition de phase . . . . .	224
6.7.4	Estimation de la variation de la rigidité sur la séquence de <i>Poly(dA)</i> étudiée dans le chapitre précédent . . . . .	225
CONCLUSION . . . . .		226
ARTICLE . . . . .		229

## INTRODUCTION

Comprendre les effets de l'environnement direct de la molécule d'ADN sur la mécanique de cette dernière est essentiel pour mieux appréhender et comprendre ce qui se produit à l'intérieur du noyau cellulaire. Notamment pour appréhender les événements et mécanismes mis en jeu lors de processus biologiques tels que la transcription ou la réplication. En effet, il a été constaté expérimentalement que l'ADN est un biopolymère qui peut se rigidifier, se courber abruptement, voire se dénaturer en fonction de sa séquence et en fonction de la fixation de protéines sur certains sites spécifiques. La détection de tels événements locaux, qui modifient radicalement les propriétés physiques telles que la rigidité et l'élasticité de la molécule d'ADN, constitue un défi crucial.

Par exemple, la fixation d'une protéine de diamètre moyen  $\sim 5$  nm sur une séquence de quelques paires de base, de l'ordre du nanomètre, implique fort probablement de nouvelles contraintes mécaniques sur la structure de l'ADN. Afin de contenir et relaxer ces contraintes, apparaît sur le polymère une variation et/ou modification structurale qui influe sur sa conformation à plus ou moins grande échelle. Intéressons-nous, par exemple, au phénomène d'allostérie, historiquement observé sur les protéines, et depuis quelques décennies étudié sur la molécule d'ADN. Un site allostérique de l'ADN interagit avec une molécule tierce, ceci peut provoquer une modification des propriétés structurales du bio-polymère et ainsi, activer ou faciliter sa capacité d'interaction avec d'autres acteurs biologiques. Il a été démontré que l'activation de sites allostériques de l'ADN peut avoir des conséquences fonctionnelles à longue distance lors du phénomène de compaction de l'ADN sous sa forme de chromatine [LFC<sup>+</sup>15].

Les implications de la fixation des protéines sur l'ADN sont nombreuses. Tout d'abord les protéines interagissant avec l'ADN peuvent régir et réguler l'activation et le déroulement de nombreux processus biologiques. De nombreuses études ont pu démontrer que la fixation de protéines sur l'ADN peut se traduire par la formation de **courbures locales** [WC84, PDCR09]. La valeur de ces angles de **courbure extrinsèque**, car provoquée par la fixation d'une molécule tierce, peut varier d'une dizaine de degrés aux environs de  $160^\circ$ , pour notamment la fameuse *Integration Host Factor* (IHF), protéine essentielle à l'organisation du chromosome bactérien [RYMN96]. Au-delà de l'effet sur la statistique de conformation de l'ADN, des études ont mis en exergue le rôle majeur de la fixation de protéines sur la dynamique de l'ADN, notamment lors du remodelage de la chromatine [MFMC14]. L'effet induit est donc loin d'être négligeable. Nous est-il possible d'observer et de quantifier l'effet induit par la présence d'une courbure de l'ADN ?

De plus, leurs interactions s'effectuent sur des **sites de fixation spécifique**. Autrement dit, certaines protéines peuvent interagir avec l'ADN seulement sur leur site de fixation. L'événement correspondant à la fixation de cette dernière sur une séquence quelconque reste possible mais peut posséder une très faible durée de vie. Si nous nous intéressons à cet événement, il est alors possible que sa signature sur la structure et la conformation de l'ADN ne soit pas distinguable via tous les dispositifs expérimentaux. La protéine peut se *décrocher* du polymère aussi vite qu'elle s'y est fixée avec un temps caractéristique trop rapide pour notre système de détection. En revanche, certaines séquences spécifiques ont une influence majeure sur la statistique et la dynamique conformationnelle de l'ADN.

Dans le chapitre précédent, nous avons déjà pu apprécier certains effets induits par la présence d'un *Poly(dA)* de 60 bp sur le comportement global d'un ADN de 583 bp. Quelles sont les particularités de sites spécifiques courts, correspondant à une séquence de l'ordre d'une dizaine de bases ? Ces dernières influent-elles sur le comportement global de l'ADN ?

Afin d'explorer les effets dus à la présence de courbures, éléments primordiaux dans la vie de l'ADN *in vivo*, ainsi que les implications liées à la présence de sites spécifiques dans l'ADN, nous nous sommes intéressés à des **courbures locales intrinsèques**. En effet, certaines séquences bien connues, induisent la formation d'une courbure locale dans la structure même de l'ADN. Ceci s'effectue sans l'implication d'autres molécules. La caractérisation d'une **courbure locale intrinsèque**, utilisée comme **courbure modèle**, va nous permettre d'élaborer une méthode de quantification d'un angle de courbure quelconque. En effet, les variations de signal mesurées seraient ainsi seulement induites par la présence de la courbure. Aucun effet, tels que des interactions électrostatiques ou de l'encombrement stérique, ne viendra perturber et influencer la mesure de l'angle de courbure. Cette étude modèle est nécessaire avant de songer à mesurer l'influence d'une courbure extrinsèque.

Dans ce cadre-là, nous cherchons à développer des outils pour quantifier, à la fois expérimentalement et théoriquement, l'impact de modifications locales sur la mécanique d'une molécule ADN.

Quels sont les observables qui nous sont accessibles via les expériences TPM ? Quels sont donc les effets de la présence de séquences spécifiques sur les propriétés physiques de l'ADN ? L'effet est-il semblable en fonction de la séquence ? Que se passe-t-il si les séquences sont positionnées en opposition de phase afin de compenser leurs effets ? Les simulations nous permettent-elles de modéliser et prédire cette dépendance ?

Là encore, ce chapitre est développé selon le même format que précédemment, dû au fait que l'essentiel de ces résultats sont déjà publiés. Ainsi, les principaux résultats et théories associées seront précisés mais les paramètres expérimentaux, déjà renseignés dans l'article ne seront pas détaillés.



## 6.1 SÉQUENCES A-TRACT, INDUISANT UNE COURBURE INTRINSÈQUE DANS LA MOLÉCULE D'ADN

Dès les années 80, des expériences de migration de molécules d'ADN au travers un gel d'électrophorèse, ont relevé l'existence de courbures dans l'ADN qui ont été attribuées aux propriétés de séquences spécifiques nommées **A-tracts**, "*trains*" de base A [MLCE82]. Dans cette étude nous avons choisi de nous intéresser à la séquence **A-tracts** suivante : CA<sub>6</sub>CGG car elle est connue pour induire une courbure significative. Les dénominations **A-tracts** peuvent référer à un grand nombre de séquences qui correspondent aux caractéristiques suivantes : la succession d'au minimum 3 bases Adénines. Généralement ces adénines sont entourées par des bases cytosines, ceci permet une maximisation de l'effet induit. La déclinaison des **A-tracts** peut amener à une séquence de quelques A auxquels est associée une base Thymine avant d'être flanquée de cytosine. Il est aussi possible d'interposer une base Guanine avant les C. La succession de A induit une augmentation de la compaction de l'empilement des bases. La présence de C induit une inclinaison dans la double hélice afin de relaxer la contrainte d'empilement. Ainsi, l'existence de la courbure est due à la jonction entre la succession des adénines et la cytosine adjacente qui permet de recoller aux caractéristiques de l'ADN de forme B. Il existe donc un grand nombre de séquences possibles induisant des courbures de plus ou moins grande amplitude.

Ces séquences particulières tiennent un rôle important dans la régulation de la transcription [Str85]. Ceci a contribué à accroître l'intérêt porté aux séquences **A-tracts** et à leur caractérisation. Depuis, un certain nombre d'études se sont succédées afin d'évaluer l'angle de courbure intrinsèque d'une telle séquence.

### 6.1.1 Stratégie de construction des molécules courbées

Nous nous sommes donc intéressés à la séquence **A-tracts** suivante : CA<sub>6</sub>CGG.

Nous décidons de partir de la molécule  $L_{ADN} = 575$  bp, Fig. 6.1. En effet, ce choix est basé sur l'hypothèse qu'un effet local aura plus de répercussion sur les conformations du polymère dans la mesure où la taille de la séquence *perturbante* est importante relativement à la taille de l'ADN utilisé. La taille de la séquence **A-tracts** étant de 10 bp, nous avons choisi une longueur d'ADN petite par rapport à notre gamme. Un **A-tract** représente donc  $\frac{1}{50}$  de la longueur totale de la molécule.

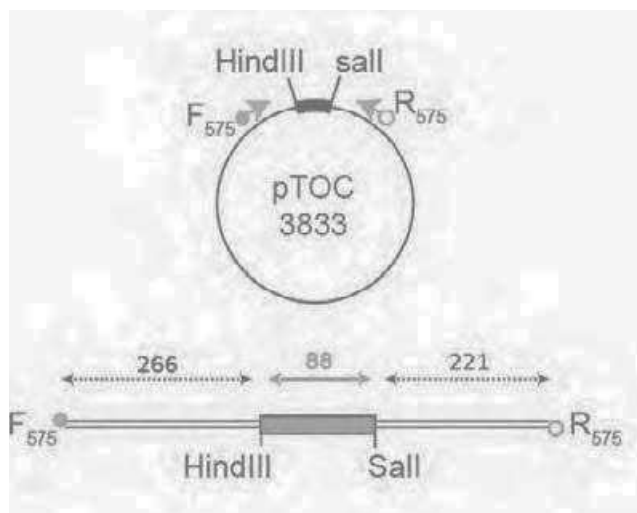


Fig. 6.1 – Schéma du plasmide et de la construction des ADN utilisés

A l'intérieur de la molécule d'ADN de 575 bp, seul un insert de 88 bp sera modifié avec une séquence variant avec le nombre de **A-tracts** que nous souhaitons introduire dans la molécule d'ADN, Fig. 6.2.

Ainsi, nous construisons plusieurs molécules d'ADN possédant un nombre croissant de **A-tracts** positionnés en phase avec le pas de l'hélice dans l'insert de 88 bp. Ceci signifie que les bases qui constituent la séquence spécifique sont positionnées et répétées de façon régulière afin de respecter le *pas* du bio-polymère de forme B qui est de 10.5 bp par tour d'hélice. Ainsi, pour être en phase la première base cytosine de la séquence CA<sub>6</sub>CGG peut être placée tous les  $n \times 10.5$  bp, avec  $n$  un entier.

La construction en phase assure, normalement, un effet additif des phénomènes de courbure. Tandis qu'une construction en opposition de phase, d'un nombre pair d'**A-tracts** doit, théoriquement, quasiment annuler les courbures deux à deux.

### 6.1.2 Détail des séquences utilisées durant cette étude

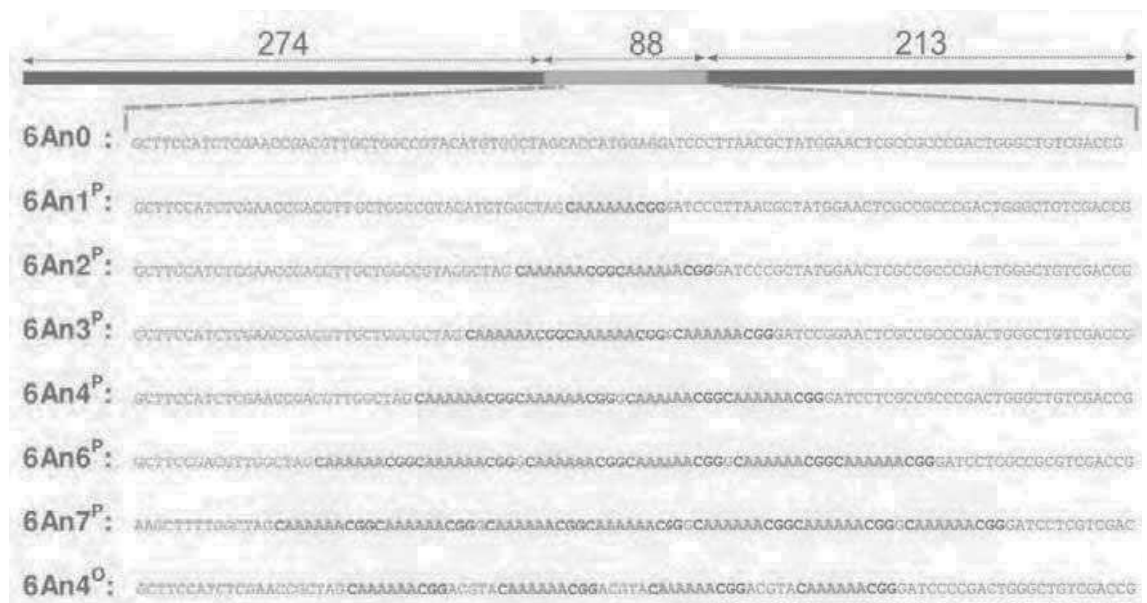


Fig. 6.2 – Détail de l'insert de 88 bp contenant les séquences courbantes. Une base G est rajoutée tous les 2 **A-tracts** afin de rattraper le pas de l'hélice pour positionner les séquences spécifiques en phase

La molécule désignée 6An0, est la molécule témoin ne possédant aucune séquence **A-tracts** dans l'insert. Les molécules 6Ani<sup>P</sup> sont constituées de  $i$  séquences **A-tracts** construites en phase, désignation **P** pour *phase*. La séquence 6An4<sup>O</sup> correspond à l'assemblage de 4 séquences **A-tracts** disposées en opposition de phase, avec **O** pour *opposition de phase*.

## 6.2 THÉORIE POUR UNE COURBURE LOCALE

Avant d'aller plus loin dans l'analyse de nos données, exprimons la dépendance théorique d'une polymère présentant une courbure locale.

### 6.2.1 Model Kinked Worm-Like-Chain

Pour cela nous considérons un polymère de longueur  $L$  décrit selon le modèle **WLC**, 2.3.2, de longueur de persistance  $L_p$  possédant un point singulier ou *kink*, à la position  $l$ . La présence du *kink* induit la formation d'une courbure spontanée dont la valeur de l'angle associée est notée  $\theta$ , Fig. 6.3.

Les caractéristiques physiques du polymère, telles que la distance bout-à-bout, sont exprimées en fonction des vecteurs tangents associés à l'abscisse curviligne  $s$ , comme exprimé en 2.21.

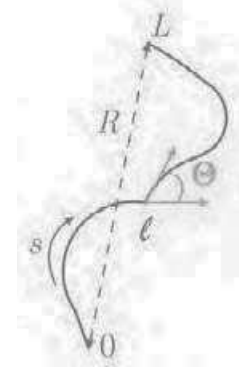


Fig. 6.3 – Représentation du polymère.

La distance bout-à-bout de l'ADN peut s'exprimer ainsi :

$$\begin{aligned}
 \langle \mathbf{R}_{\text{WLC}}^2 \rangle &= \int_0^L \int_0^L \langle \vec{t}(s) \vec{t}(s') \rangle ds ds' \\
 &= \int_0^l \int_0^l \langle \vec{t}(s) \vec{t}(s') \rangle ds ds' + 2 \int_l^L \int_0^l \langle \vec{t}(s) \vec{t}(s') \rangle ds ds' \\
 &\quad + \int_l^L \int_l^L \langle \vec{t}(s) \vec{t}(s') \rangle ds ds' \quad (6.1)
 \end{aligned}$$

Le premier et le dernier termes de l'équation 6.1 se ramènent à l'expression classique du modèle **WLC** faisant intervenir la longueur de persistance.

$$\begin{aligned}
 \langle \mathbf{R}_{\text{WLC}}^2 \rangle_{1,3} &= \int_0^l \int_0^l \exp \left[ 1 - \frac{|s' - s|}{L_p} \right] ds ds' \\
 &= 2lL_p - 2L_p^2 \left[ 1 - \exp \left( -\frac{l}{L_p} \right) \right]
 \end{aligned}$$

En utilisant les propriétés d'une chaîne de Markov, pour l'intégration double sur  $s$  et  $s'$  tels que,  $s > l > s'$ , nous pouvons exprimer le second terme de Eq. 6.1 comme :

$$2 \int_l^L \int_0^l \langle \vec{t}(s) \vec{t}(s') \rangle ds ds' = 2 \int_l^L \int_0^l \langle \vec{t}(s) \vec{t}(l^+) \rangle \cos \theta \langle \vec{t}(l^-) \vec{t}(s') \rangle ds ds'$$

Ce qui mène à l'expression suivante :

$$\begin{aligned}
 \langle \mathbf{R}_{\text{WLC}}^2 \rangle_2 &= 2 \int_l^L ds \int_0^l ds' \exp \left( -\frac{s-l}{L_p} \right) \cos \theta \exp \left( -\frac{l-s'}{L_p} \right) \\
 &= 2 \cos \theta L_p^2 \left[ 1 - \exp \left( -\frac{L-l}{L_p} \right) \right] \times \left[ 1 - \exp \left( -\frac{l}{L_p} \right) \right] \quad (6.2)
 \end{aligned}$$

Ainsi, l'expression de la distance bout-à-bout est égale à :

$$\begin{aligned}
 \langle \mathbf{R}_{\text{WLC}}^2 \rangle &= 2L_p^2 \left[ \frac{L}{L_p} - 2 + \exp \left( -\frac{l}{L_p} \right) + \exp \left( -\frac{L-l}{L_p} \right) \right] \\
 &\quad + 2L_p^2 \left[ \cos \theta \left( 1 - \exp \left( -\frac{L-l}{L_p} \right) - \exp \left( -\frac{l}{L_p} \right) + \exp \left( -\frac{L}{L_p} \right) \right) \right] \quad (6.3)
 \end{aligned}$$

### 6.2.2 Paramètres K-WLC pour une courbure au centre de la molécule

En considérant le *kink* à la position  $l = 288$  bp, soit au centre de la molécule d'ADN  $L_{ADN} = 575$  bp. L'expression 6.3 s'écrit sous la forme :

$$R_{WLC \ l=288, L_{ADN}=575} = \sqrt{\langle \mathbf{R}_{WLC}^2 \rangle_{l=288}} = D \left[ \sqrt{1 + 0.342 \times \cos\theta} \right] \quad (6.4)$$

Avec  $D$  égale à 105.3 nm, la valeur de la distance bout-à-bout pour  $\theta = \frac{\pi}{2}$ .

### 6.2.3 Paramètres K-WLC pour une courbure excentrée

En considérant maintenant le *kink* à la position  $l = 313$  bp, soit la position réelle du centre de la séquence **A-tracts** dans l'ADN  $L_{ADN} = 575$  bp utilisé durant nos expériences, l'expression 6.3 s'écrit sous la forme :

$$R_{WLC \ l=313, L_{ADN}=575} = \sqrt{\langle \mathbf{R}_{WLC}^2 \rangle_{l=313}} = D \left[ \sqrt{1 + 0.338 \times \cos\theta} \right] \quad (6.5)$$

La variation induite par la variation de 0.342 à 0.338 du préfacteur de la dépendance en  $\theta$  est juste légèrement visible pour  $\theta \rightarrow 0$  ou  $\theta \rightarrow \pi$ , comme en témoigne le graphique Fig. 6.4.

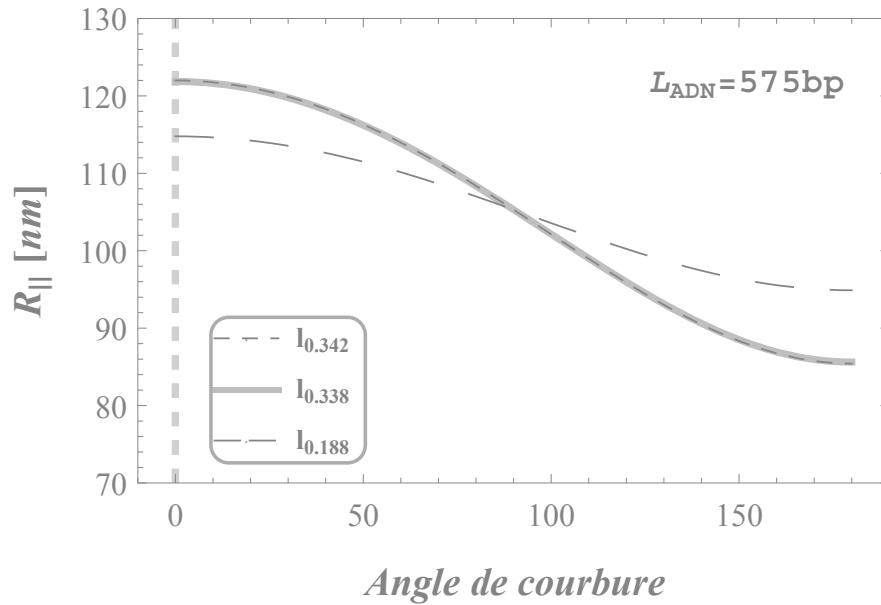


Fig. 6.4 – Variation de la distance bout-à-bout d'un polymère  $L_{ADN} = 575$  présentant un kink en position  $l = 288$  bp pour la courbe en tirets noirs qui est quasiment superposée avec la courbe verte correspondant à la position  $l = 313$  bp. La courbe en tirets longs correspond à la position du kink en  $l = 188$  bp.

Par contre, la présence d'un *kink* excentré, par exemple à  $\frac{1}{3}$  d'une extrémité de l'ADN, modifie complètement l'amplitude de l'effet induit sur  $R_{||}$  par ce dernier. Cet effet est également illustré en Fig. 6.4.

### 6.3 MESURE DE LA VARIATION DE LA DISTANCE BOUT-À-BOUT DES COMPLEXES ADN/PARTICULE EN FONCTION DE LA SÉQUENCE

Intéressons-nous à la distance bout-à-bout des complexes ADN/particule, grandeur extraite directement de nos mesures TPM et corrigée de l'effet lié au temps de corrélation du système. Observons-nous une dépendance de cette variable, observable et mesurable par TPM, en fonction du nombre de **A-tract** présente dans la molécule ?

La variation observée sur la distance bout-à-bout des complexes ADN/particule en présence de 1 à 3 **A-tract** est de l'ordre de notre erreur de mesure expérimentale. En revanche, dès la présence de 4 **A-tracts**, ou plus, la variation de  $R_{Exp||}$  est significative, allant de 149.7 à 146.5 nm et jusqu'à 140.4 nm en présence de 7 séquences spécifiques assemblées en phase, Fig. 6.5.

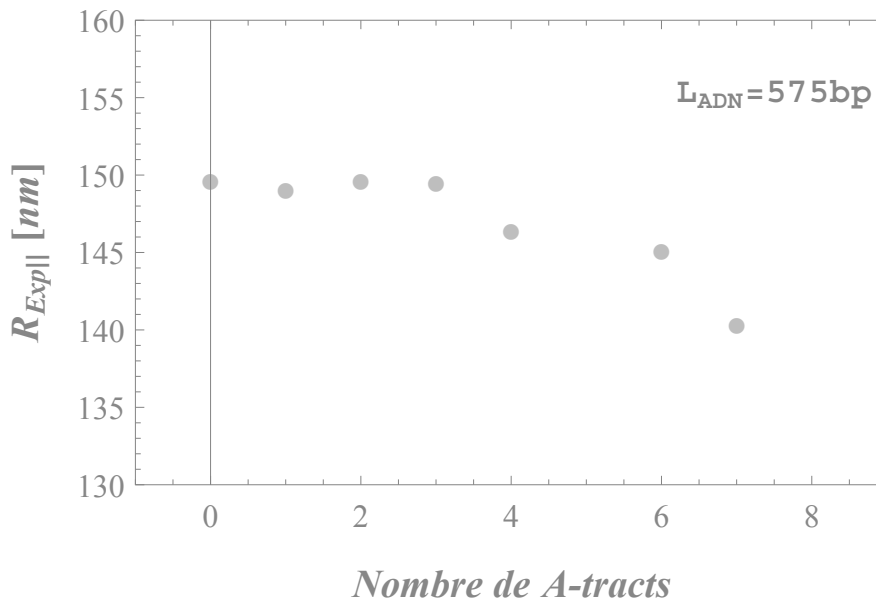


Fig. 6.5 – Variation de la distance bout-à-bout des complexes ADN/particule obtenue pour un ADN de 575 bp contenant un nombre croissant de **A-tracts**. La barre d'erreur est plus petite que la taille du symbole.

### 6.4 EXTRACTION DE LA DISTANCE BOUT-À-BOUT DE L'ADN SEUL

Une question se pose quant à la méthode de soustraction de la contribution de la particule de nos données expérimentales. En effet, nous avons déjà constaté les erreurs que peuvent introduire l'utilisation d'une méthode non adaptée. Afin de tester les biais induits par les différentes méthodes d'extraction des propriétés physiques du polymère nous expérimentons les effets de ces dernières sur nos données de simulation.

La méthode d'extraction la plus adaptée nous permet de remonter aux paramètres initiaux appliqués dans les simulations. Nous testons si l'approche *simple* ou l'approche proposée par *Segall* nous permet de respecter cette propriété.

Nous partons des amplitudes du mouvement de nos complexes ADN/particules simulées auxquelles nous appliquons la même procédure d'extraction de  $R_{ADN}$ , utilisée sur les données expérimentales. Une fois la méthode validée, ceci va nous permettre de construire une procédure simple et fiable pour quantifier la dépendance de  $R_{ADN}$  en fonction de l'angle de courbure locale.

#### 6.4.1 Simulation MC-D d'un ADN présentant une courbure locale centrée

Nous réalisons des simulations par Monte-Carlo dynamique modélisant notre système expérimental. En effet, ce modèle, donnant accès à des informations clés sur la dynamique du système, fut le premier modèle développé pour rendre compte de la géométrie associée à la méthode TPM. À l'époque de ce projet, nous disposions seulement de ce dernier.

Nous simulons la taille d'ADN suivant  $L_{ADN} = 575$  bp, en fixant le nombre de monomères à  $N = (25 + 1)$ . Afin de rendre compte de la présence des séquences **A-tracts** utilisées dans les expériences TPM, le polymère simulé présente un angle de courbure fixé entre trois monomères successifs situés au centre de la molécule d'ADN à  $l = 288$  bp. Nous imposons ainsi la présence d'une courbure locale et ponctuelle. La disposition des séquences spécifiques en une succession directe et juxtaposée de **A-tracts**, rend cette approximation valable. La gamme de valeurs pour l'angle de courbure associée est prospectée sur le domaine suivant  $\theta = 0, 18, 30, 45, 50, 60, 72, 90, 120, 180^\circ$ .

#### 6.4.2 Choix du mode d'extraction

La contribution de la particule est extraite des données des simulations via la méthode *classique* présentée en 3.8.2, et via la méthode présentée par Segall *et al.* en 2006, 3.8.2. Les données  $R_{ADN,C}$  et  $R_{ADN,S}$  sont ensuite évaluées en fonction de la qualité de leur ajustement par Eq. 6.4. Les données associées à la méthode *classique*,  $R_{ADN,C}$ , sont convenablement ajustées, contrairement aux données associées à la méthode de Segall,  $R_{ADN,S}$ , Fig. 6.6.

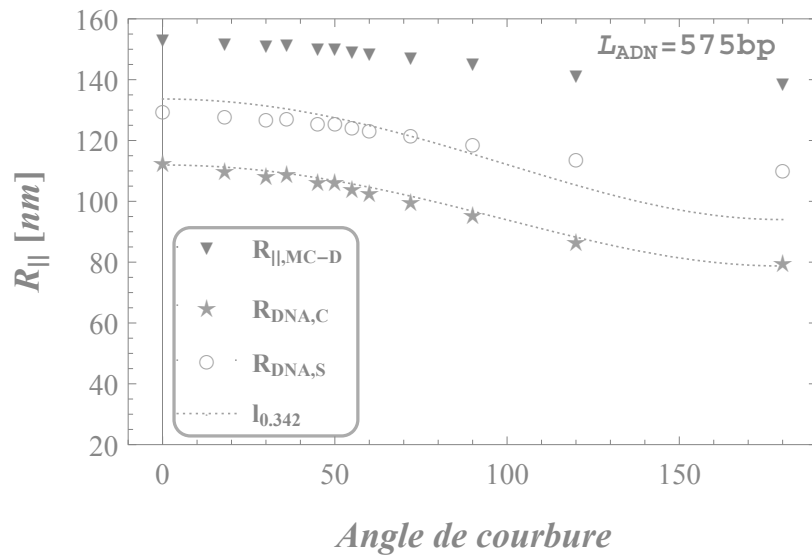


Fig. 6.6 – Variation de la distance bout-à-bout des complexes ADN/particule obtenue pour un ADN de 575 bp via des simulations numériques MC-D. Variation de  $R_{WLC}$ , extraite soit par la méthode classique soit par la méthode Segall ainsi que leur ajustement respectif via l'équation 6.4.

L'extraction de la contribution de la particule au mouvement via la méthode de Segall n'est donc pas adaptée dans le cas présent, où les conformations du polymère sont contraintes par la présence d'une courbure locale.

	Expression fit : $R_{ADN} = D \times \sqrt{1 + 0.342 \times \cos\theta}$	
	$R_{ADN,C}$	$R_{ADN,S}$
$D, \text{ nm}$	96.8	115.5

## 6.5 QUANTIFICATION DE L'ANGLE DE COURBURE À PARTIR D'UN JEU DE CONSTRUCTIONS D'ADN

La méthode *classique* est donc utilisée pour soustraire la contribution de la particule afin d'obtenir la variation de la distance bout-à-bout de l'ADN seul. Ainsi, les données pourront être analysées par l'intermédiaire d'une méthode simple et robuste permettant d'extraire les mesures de l'angle de courbure induite par la présence de **A-tracts**.

### 6.5.1 Extraction de l'angle de courbure sur un ADN supposé régulier

Afin d'évaluer la contribution d'une séquence sur l'angle de courbure, nous supposons donc que chacune des  $n$  séquences **A-tracts** successives construites en phase de l'insert, impose le même angle de courbure  $\theta_1$ . De plus, l'approximation au premier ordre donnant lieu à  $\theta = n \times \theta_1$  est validée. Ainsi les données  $R_{ADN}$  sont ajustées via l'équation suivante :

$$R_{WLC \text{ } l=313} = D \left[ \sqrt{1 + 0.338 \times \cos(n.\theta_1)} \right] \quad (6.6)$$

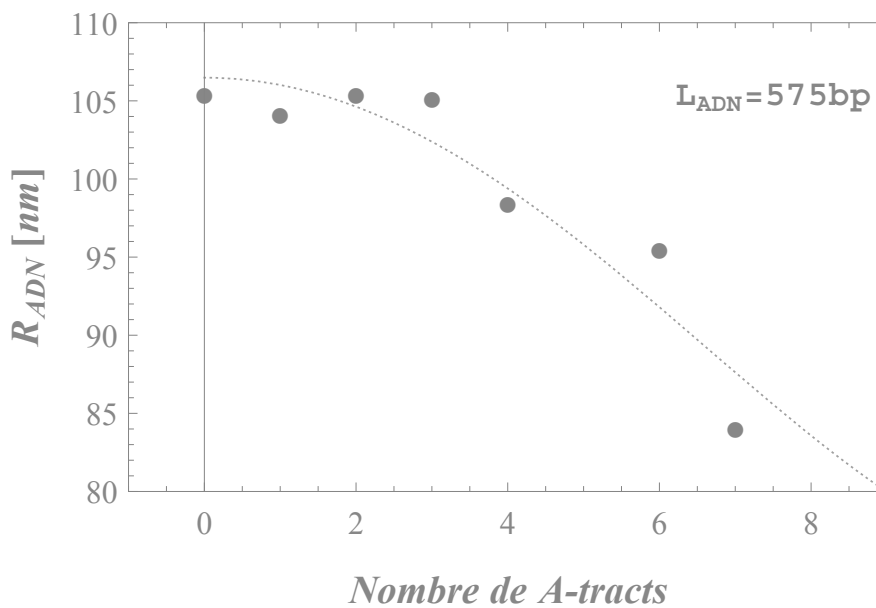


Fig. 6.7 – Variation de la distance bout-à-bout de l'ADN en fonction du nombre de **A-tracts** présent. Courbe pointillés rouges : ajustement selon Eq. 6.6, avec trois paramètres libres.

L'évaluation des données expérimentales correspond aux valeurs suivantes :

	Expression fit : $R_{ADN} = D \left[ \sqrt{1 + 0.338 \times \cos(n.\theta_1)} \right]$
$D, \text{ nm}$	$92 \pm 2$
$\theta_1$	$15^\circ \pm 2^\circ$

Cette estimation n'explique pas entièrement les dispersions observées sur nos données, Fig. 6.7. En effet, les données  $R_{ADN}$  semblent non monotones en fonction de  $n$ .

### 6.5.2 Extraction de l'angle de courbure sur un ADN supposé non-régulier

L'ajout d'une séquence CA<sub>6</sub>CGG nécessite la suppression de l'équivalent d'autant de paires de bases dans la séquence initiale. De plus, comme nous avons choisi de toujours positionner la succession des **A-tracts** centrée dans l'insert de 88 bp, les bases constituant les extrémités de celui-ci diffèrent d'un ADN à l'autre.

Il est donc possible que certains effets séquences soient supprimés ou ajoutés lors de l'ajout successif des  $n$  séquences courantes.

Un des possibles effets séquences est l'existence de courbure intrinsèque. L'ensemble de la molécule d'ADN de 575 bp est donc susceptible d'être sujet à ce type d'effets séquences. L'insertion d'un **A-tract** modifierait donc la courbure intrinsèque et globale à l'ensemble de la molécule d'ADN.

Une autre approche afin d'analyser ces données est de supposer l'existence d'une courbure intrinsèque et globale à l'ensemble de la molécule d'ADN initial 6An0 regroupée sous le terme  $\theta_0$ , auquel s'ajoute ensuite la courbure intrinsèque locale associée à la présence des  $n$  **A-tracts**.

La mise en équation de cette nouvelle hypothèse se traduit via l'expression de l'évaluation suivante :

$$R_{ADN,WLC \text{ } l=313} = D \left[ \sqrt{1 + 0.338 \times \cos(n.\theta_1 - \theta_0)} \right] \quad (6.7)$$

Cette approche permet une meilleure prise en compte de la dispersion observée sur nos données expérimentales, Fig. 6.8.

	Expression fit : $R_{ADN} = D \left[ \sqrt{1 + 0.338 \times \cos(n.\theta_1 - \theta_0)} \right]$
$D', \text{ nm}$	$91 \pm 2$
$\theta'_1$	$19^\circ \pm 4^\circ$
$\theta_0$	$25^\circ \pm 19^\circ$



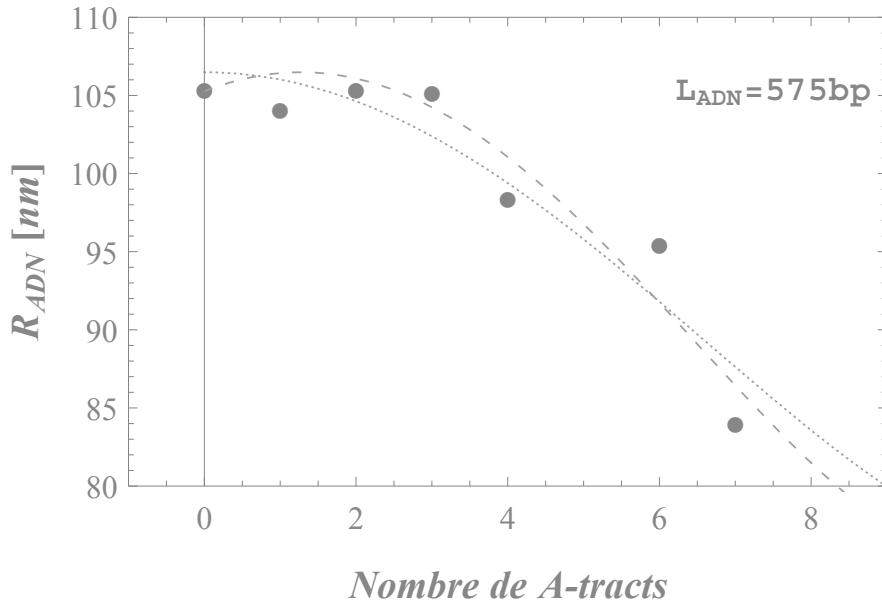


Fig. 6.8 – Variation de la distance bout-à-bout de l'ADN en fonction du nombre de A-tracts présents. Courbe pointillés rouges : ajustement selon Eq. 6.7, avec deux paramètres libres.

Cette analyse nous amène à considérer toute molécule d'ADN, loin de l'image statique d'une chaîne régulière, comme un polymère possédant une courbure spontanée globale non nulle.

### 6.5.3 Prise en compte de l'incertitude due au rayon de la particule

L'extraction de la contribution de la particule est une étape cruciale, comme nous l'avons vu. Or, le rayon de la particule ne peut être connu avec exactitude. En effet, la synthèse de cette dernière assure un rayon avec une précision de l'ordre du nanomètre. De plus, la nanoparticule est fonctionnalisée, ce qui tend à augmenter légèrement son rayon.

Afin de prendre en compte cette incertitude nous avons procédé à la même analyse que ci-dessus, lorsque la contribution de la particule a été soustraite via un rayon de  $R_{\text{par}} = 155$  nm. Se référer à l'article pour plus de détails.

La mesure finale de l'angle de courbure que nous donnons, prend en compte l'erreur induite par la mauvaise connaissance du rayon de la particule via les estimations obtenues pour  $R_{\text{par}} = 150$  nm et  $R_{\text{par}} = 155$  nm et leurs barres d'erreur respectives.

### 6.5.4 Comparaison avec les données de la littérature

La valeur de l'angle de courbure induit par la présence de séquences CA<sub>6</sub>CGG en prenant en compte l'existence d'une courbure globale le long de l'ensemble de la molécule d'ADN en considérant l'incertitude induite par l'erreur sur la mesure du rayon de la particule, est égale à  $\theta_1 = 19 \pm 4$  avec  $D = 92 \pm 2$  nm,  $\theta_0 = 25 \pm 20$ .

Cette valeur est en accord avec les principales données de la littérature utilisant des méthodes de mesures non perturbatives. Là aussi, se référer à l'article pour plus de détails.

## 6.6 PROPOSITION D'UNE FORMULE SIMPLE AFIN D'ÉVALUER L'ANGLE DE COURBURE À PARTIR DE DEUX MESURES

La stratégie présentée ci-dessus permet d'extraire la valeur d'un angle de courbure induite par une séquence spécifique avec une bonne précision. Cependant, cette approche nécessite la construction d'un certain nombre de molécules d'ADN possédant cette séquence spécifique, répétée en phase. Ceci nécessite donc un travail assez lourd.

Afin de fournir une procédure d'extraction d'un angle de courbure locale plus directe, nous tirons de notre modèle K-WLC Eq. 6.3, une formule d'interpolation simple. Cette dernière nous permet d'extraire la valeur d'un angle  $\theta_1$  en nécessitant seulement deux mesures de la distance bout-à-bout de l'ADN avec,  $R_{ADN}(\theta)$ , et sans,  $R_{ADN}(0)$ , la présence de la courbure.

$$\cos \theta = 1 + \frac{\exp\left(-\frac{L}{L_p}\right) + \frac{L}{L_p} - 1}{1 - \exp\left(-\frac{L-l}{L_p}\right) - \exp\left(-\frac{l}{L_p}\right) + \exp\left(-\frac{L}{L_p}\right)} \times \left[ \frac{R_{ADN}^2(\theta) - R_{ADN}^2(0)}{R_{ADN}^2(0)} \right] \quad (6.8)$$

### 6.6.1 Extraction de l'angle de courbure de données de la littérature

Cette formule nous permet par ailleurs d'étendre l'application aux courbures extrinsèques induites par la fixation de protéines sur l'ADN, comme motivé dans l'introduction de ce chapitre.

L'application avec succès de cette méthode d'interpolation est détaillée dans l'article.

### 6.6.2 Extraction de l'angle des courbures de la séquence TA<sub>3</sub>T

L'un de nos ADN de 575 bp possède trois **A-tracts** en phase, Fig. 6.9. Ces **A-tracts** sont constitués de seulement 3 adénines successives, mais ces derniers sont espacés les uns des autres de 30 bp contrairement aux constructions précédentes, qui les juxtaposaient.

#### Détails des séquences utilisées durant cette étude

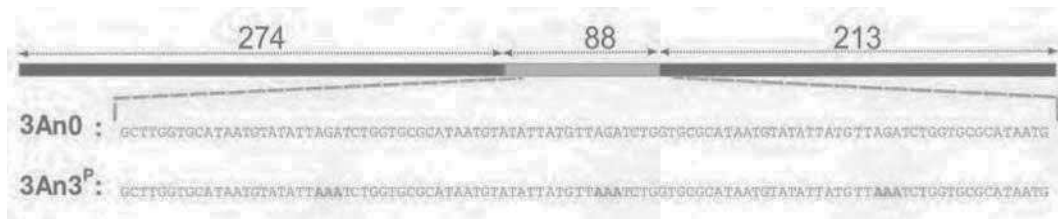


Fig. 6.9 – Détail de l'insert de 88 bp contenant les séquences courbantes

Cette séquence **A-tracts** possède ses extrémités flanquées de bases thymines. En théorie, pour un **A-tracts**, cette particularité induit une courbure moins importante que si les extrémités correspondaient à des bases cytosines ou guanines [WC84].

### Extraction de la distance bout-a-bout de l'ADN seul

La méthode *classique* de la physique des polymère est utilisée afin d'extraire  $R_{ADN}$  des données  $R_{Exp||}$ .

Séquence	$R_{Exp  }$	$R_{ADN}$
Sans TA <sub>3</sub> T	$151.7 \pm 1.0$	$109.6 \pm 0.5$
Avec TA <sub>3</sub> T	$148.0 \pm 0.3$	$101.8 \pm 0.9$

### Estimation de la courbure induite par un TA<sub>3</sub>T

La séquence spécifique est centrée sur la position  $l = 313$  bp dans l'ADN  $L_{ADN} = 575$  bp.

L'utilisation de la formule d'extrapolation, Eq. 6.8, mène à l'estimation d'un angle total de courbure de  $63.2^\circ \pm 7^\circ$  soit un angle de  $21.2^\circ \pm 2^\circ$ , induit par chaque séquence TA<sub>3</sub>T.

La courbure induite est plus importante que pour les CA<sub>6</sub>CGG. Notre méthode d'estimation de l'angle de courbure repose sur les hypothèses que la courbure est locale et **ponctuelle**. Cependant, dans le cas de l'ADN présentant 3 séquences TA<sub>3</sub>T, cette dernière hypothèse n'est pas entièrement justifiée. En effet, les séquences courbantes commencent à être éloignées les unes des autres. Ceci peut induire une erreur dans l'estimation de la valeur de l'angle de courbure associé.

### 6.6.3 Extraction de l'angle de courbure de séquences A-tract variées

D'autres types de séquences **A-tract** construits en phase et inclus dans l'insert de 88 bp de l'ADN 575 bp sont à notre disposition, Fig. 6.10. Ces ADN possèdent plusieurs natures de **A-tracts** dans un même insert.

#### Détails des séquences utilisées durant cette étude

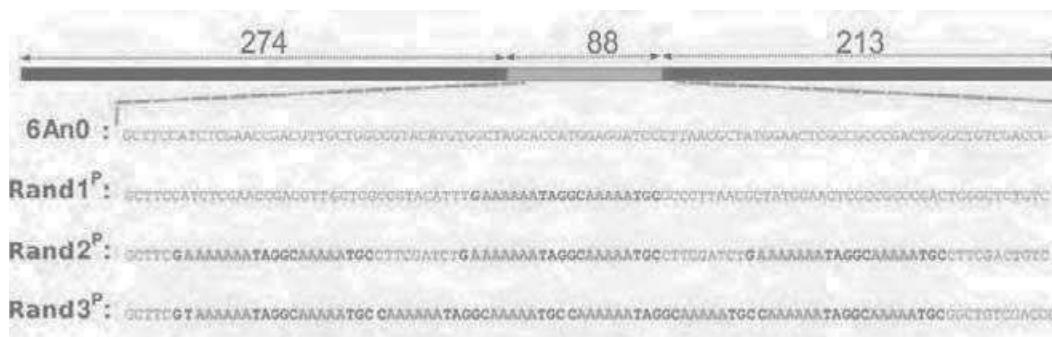


Fig. 6.10 – Détail de l'insert de 88 bp contenant les séquences courbantes aléatoire

La séquence 6An0, molécules d'ADN ne possédant aucun **A-tract**, est utilisée comme séquence de référence comme précédemment.

### Extraction de la distance bout-à-bout de l'ADN seul

La méthode classique est utilisée afin d'extraire  $R_{ADN}$  des données  $R_{Exp||}$ .

Séquence	$R_{Exp  }$	$R_{ADN}$
6An0	$149.7 \pm 0.2$	$105.4 \pm 0.4$
Rand1 <sup>P</sup>	$148.3 \pm 0.3$	$102.4 \pm 0.9$
Rand2 <sup>P</sup>	$143.5 \pm 0.4$	$91.6 \pm 1.1$
Rand3 <sup>P</sup>	$141.2 \pm 0.2$	$86.1 \pm 0.6$

### Estimation de la courbure induite par ces séquences

La séquence spécifique est à la position  $l = 313$  bp dans l'ADN  $L_{ADN} = 575$  bp.

L'utilisation de la formule d'extrapolation, Eq. 6.8, mène à l'estimation d'un angle de courbure total présent dans l'insert.

- Pour l'ADN Rand1<sup>P</sup>, l'angle de courbure total est estimé à  $39^\circ \pm 8^\circ$ .
- Pour l'ADN Rand2<sup>P</sup>, l'angle de courbure total est estimé à de  $88^\circ \pm 8^\circ$  soit un angle moyen induit par la juxtaposition des 2 **A-tracts** répétés 3 fois en phase, de  $29^\circ \pm 3^\circ$ .
- Pour l'ADN Rand3<sup>P</sup>, l'angle de courbure total est estimé à  $109^\circ \pm 5^\circ$ .

L'augmentation de la valeur estimée de l'angle de courbure total induit dans la molécule d'ADN est cohérente avec l'augmentation du nombre de séquences courbantes présentes.

## 6.7 THÉORIE POUR UN INSERT RIGIDE LOCAL

### 6.7.1 Modèle *Stiffer Worm-Like-Chain*

Une autre approche à discuter correspond à considérer la séquence spécifique comme une séquence induisant une variation locale de la rigidité, plus précisément de la longueur de persistance de la molécule d'ADN.

Soit un polymère de longueur  $L$  décrit selon le modèle **WLC**, 2.3.2, de longueur de persistance  $L_p$  possédant un insert de rigidité différente à la position  $l$  et de longueur  $\alpha$ , Fig. 6.11.

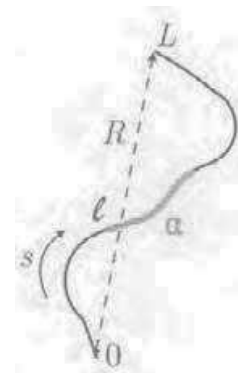


Fig. 6.11 – Représentation du polymère discuté dans le texte

Les caractéristiques physiques du polymère, telle que la distance bout-à-bout, sont exprimées en fonction des vecteurs tangents associés à l'abscisse curviligne  $s$ , comme exprimé en 2.21.

La distance bout-à-bout de l'ADN peut s'exprimer ainsi :

$$\begin{aligned}
\langle \mathbf{R}'_{\text{WLC}} \rangle^2 &= \int_0^L \int_0^L \langle \vec{t}(s) \vec{t}(s') \rangle ds ds' \\
&= \int_0^l \int_0^l \langle \vec{t}(s) \vec{t}(s') \rangle ds ds' + \int_l^{l+\alpha} \int_l^{l+\alpha} \langle \vec{t}(s) \vec{t}(s') \rangle ds ds' \\
&+ \int_{l+\alpha}^L \int_{l+\alpha}^L \langle \vec{t}(s) \vec{t}(s') \rangle ds ds' + 2 \int_0^l \int_l^{l+\alpha} \langle \vec{t}(s) \vec{t}(s') \rangle ds ds' \\
&+ 2 \int_0^l \int_{l+\alpha}^L \langle \vec{t}(s) \vec{t}(s') \rangle ds ds' + 2 \int_l^L \int_l^L \langle \vec{t}(s) \vec{t}(s') \rangle ds ds' \quad (6.9)
\end{aligned}$$

Comme lors du développement précédent, ce calcul mène à l'expression suivante :

$$\begin{aligned}
\langle \mathbf{R}'_{\text{WLC}} \rangle^2 &= 2L_p^2 \left[ \frac{L-\alpha}{L_p} + \exp\left(-\frac{l}{L_p}\right) + \exp\left(-\frac{L-\alpha-l}{L_p}\right) - 2 \right. \\
&+ \left. \exp\left(-\frac{\alpha}{L_{p,\alpha}}\right) \times \left(1 - \exp\left(-\frac{l}{L_p}\right)\right) \left(1 - \exp\left(-\frac{L-\alpha-l}{L_p}\right)\right) \right] \\
&+ 2L_p L_{p,\alpha} \left(1 - \exp\left(-\frac{\alpha}{L_{p,\alpha}}\right)\right) \left(2 - \exp\left(-\frac{l}{L_p}\right) - \exp\left(-\frac{L-\alpha-l}{L_p}\right)\right) \\
&+ 2L_{p,\alpha}^2 \left(\frac{\alpha}{L_{p,\alpha}} - 1 + \exp\left(-\frac{\alpha}{L_{p,\alpha}}\right)\right) \quad (6.10)
\end{aligned}$$

### 6.7.2 Argument justifiant cette approche

Théoriquement, l'ADN construit avec quatre **A-tracts** en opposition de phase devrait voir les courbures induites par les séquences spécifiques se compenser deux à deux. Nous nous attendons donc à mesurer une distance bout-à-bout pour cet ADN 6An4<sup>O</sup> proche de celle de l'ADN 6An0.

Séquence	$R_{Exp  }$	$R_{ADN}$	$l$	$\alpha$
6An0	$149.7 \pm 0.2$	$105.4 \pm 0.4$	-	-
6An4 <sup>P</sup>	$146.5 \pm 0.2$	$98.5 \pm 0.6$	298 bp	41 bp
6An4 <sup>O</sup>	$151.2 \pm 0.6$	$108.6 \pm 0.2$	291 bp	55 bp

La mesure associée à l'ADN construit en opposition de phase est comme attendu, supérieure à celle de l'ADN possédant 4 **A-tracts** en phase, que se soit sur la mesure de  $R_{Exp||}$ , Fig. 6.12 ou de  $R_{ADN}$ , Fig. 6.13. En revanche, de façon inattendue cette mesure est supérieure, l'écart est supérieur à la barre d'erreur de façon significative, à la valeur associée à l'ADN témoin 6An0, ne possédant aucune courbure.

Ceci laisse supposer que la présence des **A-tracts** induit un effet supplémentaire à la formation d'une courbure. Une augmentation de rigidité locale, pourrait expliquer une augmentation de la distance bout-à-bout de l'ADN en présence de ces séquences spécifiques.

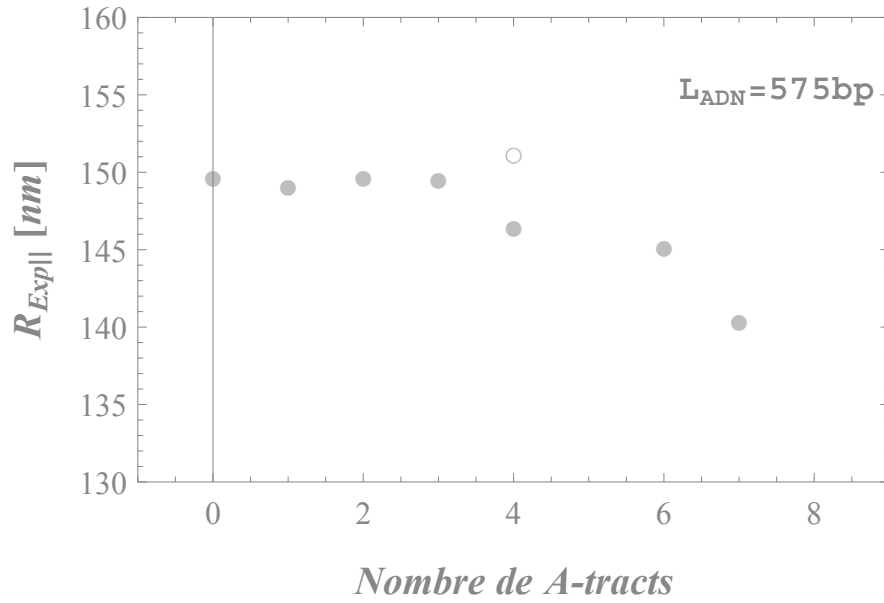


Fig. 6.12 – Variation de la distance bout-à-bout des complexes ADN/particule obtenue. Le symbole vide correspond à l'ADN 6An4<sup>O</sup>, soit 4 séquences **A-tracts** en opposition de phase selon le pas de l'hélice. La barre d'erreur est plus petite que la taille du symbole. (Constructions détaillées en Fig. 6.2)

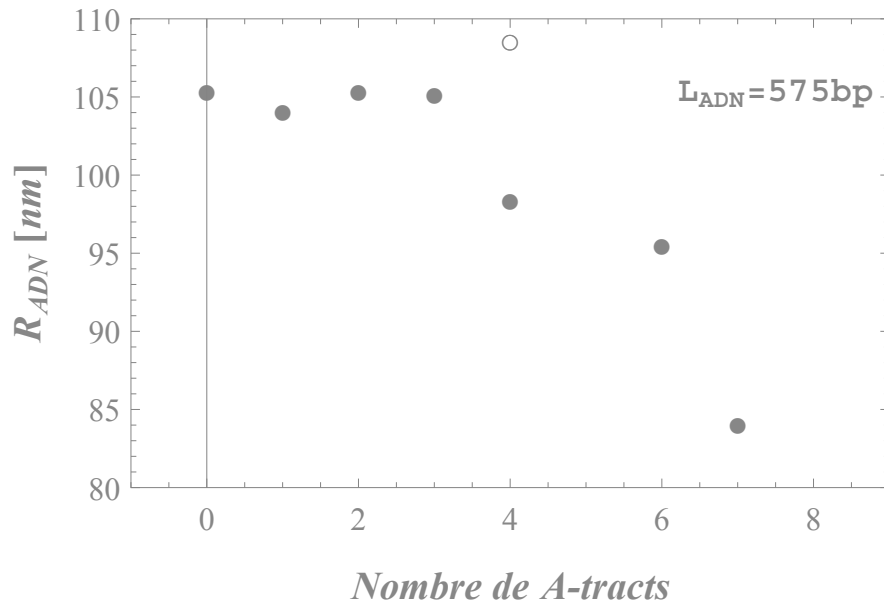


Fig. 6.13 – Variation de la distance bout-à-bout de ADN seul obtenue pour un ADN de 575 bp contenant un nombre croissant de **A-tracts**. Le symbole vide correspond à l'ADN 6An4<sup>O</sup>, soit 4 séquences **A-tracts** en opposition de phase selon le pas de l'hélice. La barre d'erreur est plus petite que la taille des symboles.

### 6.7.3 Estimation de la variation de la rigidité sur la séquence de **A-tracts** en opposition de phase

La résolution de l'équation exacte 6.10, pour l'ADN 6An4<sup>O</sup> possédant  $\alpha = 55$  bp modifié par l'insertion de séquence spécifique à la position  $l = 291$  bp et dont la longueur de persistance pour le reste de la séquence est supposée égale à  $L_p[6AnO] = 34.5$  nm, extraite via le modèle WLC, conduit à l'estimation  $L_{p,\alpha} = 71.7 \pm 9$  nm pour la longueur de persistance associée aux quatre **A-tract** en opposition de phase. Ceci correspond à une augmentation relative de la longueur

de persistance, associée à cette succession de séquences de  $\alpha = 55$  bp avec et sans la présence de **A-tracts** en opposition de phase, proche de 100%.

D'autres approches permettant d'estimer la variation de la longueur de persistance associée à la présence de **A-tract** sont expliquées et détaillées dans le papier.

#### 6.7.4 Estimation de la variation de la rigidité sur la séquence de *Poly(dA)* étudiée dans le chapitre précédent

La présence de *Poly(dA)* induit un effet qui tend à diminuer la distance bout-à-bout de l'ADN en présence de cette séquence spécifique. Autrement dit, dans ce cas nous cherchons l'estimation de la diminution de la rigidité locale de l'ADN.

Pour cela nous nous intéressons aux mesures réalisées dans les conditions salines physiologiques, soit  $C_{ion,Na^+} = 155$  mmol/L ou  $I = 165.1$  L/mol, comparables aux données de ce chapitre.

Séquence	$R_{Exp  }$	$R_{ADN}$	$l$	$\alpha$
Sans <i>Poly(dA)</i>	$151.2 \pm 0.4$	$108.6 \pm 1.7$	-	-
Avec <i>Poly(dA)</i>	$145.0 \pm 0.6$	$96.4 \pm 2.3$	298 bp	50 à 60 bp

La résolution de l'équation exacte Eq. 6.10, pour l'ADN incluant un *Poly(dA)* de  $\alpha = 60$  bp à la position  $l = 298$  bp et dont la longueur de persistance pour le reste de la séquence est supposée égale à celle de l'ADN sans *Poly(dA)*, supposée égale à  $L_p = 36.4$  nm via l'utilisation du modèle WLC, conduit à l'estimation  $L_{p,\alpha} = 7 \pm 6$  nm pour la longueur de persistance associée à la séquence *Poly(dA)*. Soit une diminution relative de la longueur de persistance, associée à la présence du *Poly(dA)* de l'ordre de 80%.

En utilisant l'approche permettant de rendre compte de l'effet induit par la présence d'une courbure locale spontanée, nous évaluons l'angle de courbure pouvant être associé à la présence du *Poly(dA)*. L'utilisation de la formule d'extrapolation, Eq. 6.8, mène à l'estimation d'un angle de courbure totale de  $81.0^\circ \pm 15^\circ$  induit par la séquence spécifique de 60 bp de *Poly(dA)*.

La remarque principale reste que notre méthode d'estimation, associée à une variation de la rigidité, est peu précise sur l'estimation d'une augmentation de cette dernière. Une question se pose quant au réalisme de cette extraction. Afin de valider ou invalider cette approche de plus amples mesures expérimentales et travaux semblent nécessaires. De plus, l'approche HT-TPM combinée au modèle K-WLC ne nous permet pas de discriminer facilement une augmentation de la rigidité induite par les inserts de celle due à une courbure intrinsèque.

## CONCLUSION DU CHAPITRE

Comprendre et mieux appréhender les effets qui se produisent à l'intérieur du noyau cellulaire et lors des processus biologiques passe aussi par la caractérisation d'effets pouvant être induits par la présence de protéines. Une conséquence des plus répandue découlant de la fixation de protéine est la courbure de la molécule d'ADN.

Afin de découpler les effets liées à la fixation d'une protéine et celles liées à la présence d'une courbure, nous avons décidé d'étudier les propriétés d'une courbure d'ADN modèle. Pour cela nous nous sommes intéressés aux répercussions engendrées par la présence d'une **courbure locale et intrinsèque** à la séquence-même de la molécule d'ADN, sur les propriétés physiques et mécaniques de cette dernière.

Initialement étudiées *in vivo* dans l'ADN de certains parasites, les séquences spécifiques **A-tracts** sont connues pour induire la formation d'une **courbure locale et intrinsèque**. La présence de courbure se répercute sur la longueur apparente de la molécule. Là encore, la technique TPM de molécule unique est donc idéale et adaptée pour sonder ce type de spécificité, sans contraindre significativement les degrés de libertés de la molécule.

La séquence **A-tracts** étudiée est constituée de la successions des bases azotées suivantes : CA<sub>6</sub>CGG. Cette dernière est l'une des séquences **A-tracts**, impliquant l'une des plus grandes amplitudes de courbure. De plus, afin d'obtenir une meilleure précision de mesure, la stratégie adoptée est la construction et l'élaboration d'un ensemble de molécules d'ADN incluant un nombre croissant de séquences CA<sub>6</sub>CGG. Ces séquences sont ajoutées en phase, selon le pas de l'hélice, afin d'assurer une additivité des effets induisant une courbure.

Tout d'abord, ceci nous a permis de relever une forte dépendance de la distance bout-à-bout de nos complexes ADN/particule en fonction du nombre de séquences **A-tracts** présentes dans le polymère.

En nous appuyant sur la théorie associée au modèle **WLC**, nous avons explicité un modèle **Kinked-WLC**. Ce modèle prend en compte la présence d'une courbure **ponctuelle** à la position  $l$  sur la variation de la distance bout-à-bout de l'ADN seul  $R_{ADN}$ . Ceci nous a permis de jauger les méthodes d'extraction de ce paramètre afin d'évaluer la plus adaptée à ce cas précis. En effet, les résultats numériques de notre modèle de simulation Monte-Carlo dynamique ont été appréhendés selon les deux méthodes d'extraction de la contribution de la particule au mouvement des complexes ADN/particules. Cette analyse a permis de mettre en exergue que le modèle *classique* est mieux adapté.

La dépendance observée sur la mesure de  $R_{Exp||}$  se répercute, comme attendue, sur les propriétés de l'ADN seul. Lors de l'extraction de  $R_{ADN}$  de nos mesures expérimentales, nous avons estimé l'erreur induite par l'incertitude sur le rayon de la particule lors de la soustraction de sa contribution. La valeur ainsi obtenue pour l'angle de courbure induite par la présence d'une séquence CA<sub>6</sub>CGG est de  $16^\circ \pm 2^\circ$ .



Or, la dépendance de la grandeur  $R_{ADN}$  en fonction du nombre de séquences spécifiques CA<sub>6</sub>CGG présentes dans la molécule d'ADN est non monotone. Ceci laisse supposer que notre méthode de construction d'une série de molécules courbées influe sur les particularités intrinsèques du reste de la molécule. L'hypothèse formulée conduit à considérer toute molécule d'ADN comme possédant une courbure spontanée globale non nulle, contrairement à l'image statique d'une molécule essentiellement régulière, qui peut lui être attribuée. Le modèle, ainsi amélioré par la prise en compte d'une courbure globale propre à la molécule d'ADN, permet une évaluation plus fidèle des données expérimentales. Ceci souligne la véracité de cette hypothèse. La valeur ainsi obtenue pour l'angle de courbure induit par la présence d'une séquence CA<sub>6</sub>CGG, est de  $19^\circ \pm 4^\circ$ . La valeur associée à la courbure globale de la molécule d'ADN est estimée égale à  $25^\circ \pm 19^\circ$ . Cette valeur est obtenue avec une faible précision, à cause de la suppression de séquences différentes pour chaque molécule constituant notre jeu d'ADN ainsi que l'incertitude sur  $R_{par}$ .

Forts des études précédentes sur cette séquence spécifique, nous avons pu comparer notre estimation et valider notre méthode de mesure et d'analyse.

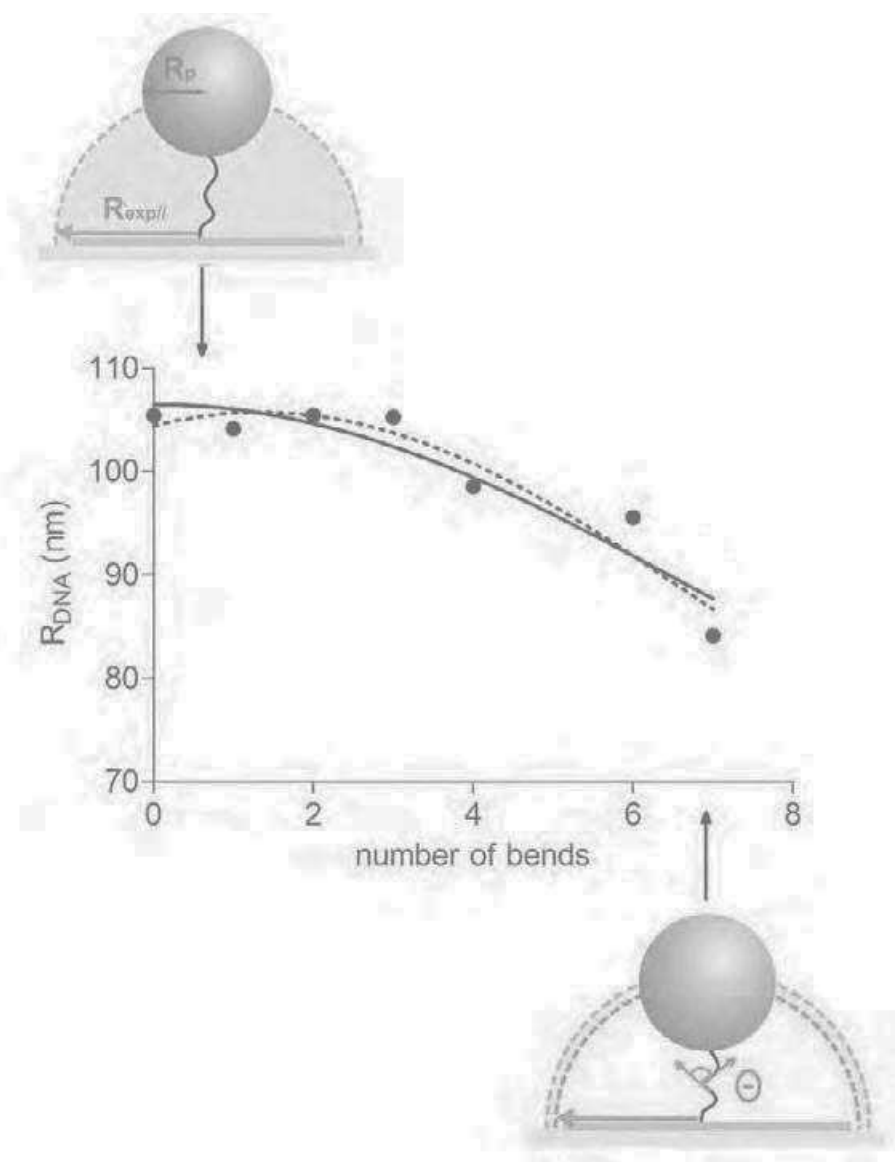
L'élaboration d'une formule simple nous a permis d'estimer les courbures induites par d'autres objets biologiques. L'utilisation de cette dernière repose sur la connaissance de seulement 2 mesures. En effet, elle nécessite seulement une mesure de la distance bout-à-bout de l'ADN seul ainsi que celle pour l'ADN incluant la séquence spécifique à étudier avec ou sans protéine. La validité de cette formule est certifiée par l'estimation de l'angle de courbure induit par la séquence CA<sub>6</sub>CGG utilisée, sans la mise en relief apportée par l'analyse globale présentée plus haut sur l'ensemble du jeu de molécules courbées. L'estimation de l'angle de courbure induit par la séquence CA<sub>6</sub>CGG est alors de l'ordre de  $15^\circ \pm 2^\circ$ . Ainsi, nous avons pu appliquer cette nouvelle approche aux autres données de la littérature. Ceci nous a permis de notamment évaluer l'angle de courbure induit par la présence de la protéine IHF. Les mesures de cristallographie indiquent une valeur d'angle de l'ordre de  $160^\circ$ , ce qui est en accord avec notre évaluation de  $180^\circ$ . Soulignons que la distance bout-à-bout tend à saturer pour les forts angles de courbures et que la méthode est alors moins précise.

Mise en lumière par l'augmentation significative de  $R_{Exp||}$  et  $R_{ADN}$  lorsque la molécule présente quatre séquences **A-tracts** en opposition de phase, une autre approche est envisageable. Cette dernière approche repose sur la prise en compte d'une variation **locale et ponctuelle** de la longueur de persistance de l'ADN. La présence d'une séquence spécifique peut conduire à une modification de la rigidité du polymère. La prise en compte conjointe d'une courbure **locale et ponctuelle** combinée à une modification **locale et ponctuelle** de la rigidité du polymère peut être explorée, détail présent dans l'article. Cette approche, combinant les deux effets, conduit à l'estimation de l'angle de courbure de  $19.5^\circ \pm 14^\circ$  et d'un terme de rigidité, défini en donnée supplémentaire S<sub>3</sub> de l'article, égal à  $2 \pm 4$  nm le tout associé à la présence d'une séquence spécifique CA<sub>6</sub>CGG. Cette estimation est réalisée dans la limite où la molécule d'ADN possède une courbure globale,  $\theta_0$ , nulle.

En nous appuyant, maintenant sur l'approche considérant seulement une modification locale du paramètre régissant la rigidité et la flexibilité de l'ADN, nous avons pu revenir sur les mesures de l'ADN présentant une séquence *Poly(dA)*. La présence d'une séquence, comme souligné dans le chapitre précédent, induit un effet sur les propriétés physiques de l'ADN. Ainsi, la prise en compte d'une modification **locale et ponctuelle** de la rigidité du polymère conduit à l'estimation d'une diminution de la rigidité de l'ordre de 80% associée à la présence de 60 bp de cette séquence spécifique *Poly(dA)*.

Ainsi, il est possible d'aller plus loin dans l'étude des effets locaux intervenant lors de processus biologiques.

ARTICLE : *Probing a label-free local bend in DNA by single molecule tethered particle motion*



# Probing a label-free local bend in DNA by single molecule tethered particle motion

Annaël Brunet<sup>1,2,3,4,†</sup>, Sébastien Chevalier<sup>3,4,†</sup>, Nicolas Destainville<sup>1,2</sup>, Manoel Manghi<sup>1,2</sup>, Philippe Rousseau<sup>5,6</sup>, Maya Salhi<sup>5,6</sup>, Laurence Salomé<sup>3,4</sup> and Catherine Tardin<sup>3,4,\*</sup>

<sup>1</sup>CNRS; LPT (Laboratoire de Physique Théorique); UMR UPS-CNRS 5152; 118 route de Narbonne, F-31062 Toulouse, France, <sup>2</sup>Université de Toulouse; UPS; LPT; F-31062 Toulouse, France, <sup>3</sup>CNRS; IPBS (Institut de Pharmacologie et de Biologie Structurale); UMR UPS-CNRS 5089; 205 route de Narbonne, F-31077 Toulouse, France, <sup>4</sup>Université de Toulouse; UPS; IPBS; F-31077 Toulouse, France, <sup>5</sup>CNRS; LMGM (Laboratoire de Microbiologie et Génétique Moléculaires); UMR CNRS-UPS 5100; 118 route de Narbonne, F-31062 Toulouse, France and <sup>6</sup>Université de Toulouse; UPS; LMGM; F-31062 Toulouse, France

Received January 6, 2015; Revised February 11, 2015; Accepted March 1, 2015

## ABSTRACT

Being capable of characterizing DNA local bending is essential to understand thoroughly many biological processes because they involve a local bending of the double helix axis, either intrinsic to the sequence or induced by the binding of proteins. Developing a method to measure DNA bend angles that does not perturb the conformation of the DNA itself or the DNA-protein complex is a challenging task. Here, we propose a joint theory-experiment high-throughput approach to rigorously measure such bend angles using the Tethered Particle Motion (TPM) technique. By carefully modeling the TPM geometry, we propose a simple formula based on a kinked Worm-Like Chain model to extract the bend angle from TPM measurements. Using constructs made of 575 base-pair DNAs with in-phase assemblies of one to seven 6A-tracts, we find that the sequence CA<sub>6</sub>CGG induces a bend angle of  $19^\circ \pm 4^\circ$ . Our method is successfully compared to more theoretically complex or experimentally invasive ones such as cyclization, NMR, FRET or AFM. We further apply our procedure to TPM measurements from the literature and demonstrate that the angles of bends induced by proteins, such as Integration Host Factor (IHF) can be reliably evaluated as well.

## INTRODUCTION

DNA bending was first revealed in the mid-80s on the mitochondrial DNA of trypanosomatid parasites, the kinetoplast DNA (kDNA) (1) and attributed to the intrinsic bending property of the A-tracts sequences present in kDNA

(2,3). These A-tracts were not only abundantly found in other prokaryotic and eukaryotic organisms but they were also shown to have a biological role, for example, by participating in the regulation of transcription (4–13). The binding of protein to DNA, that occurs in most of DNA-related biological processes, was also observed to induce the local bending of DNA (3,14–16).

As a result, for both intrinsic, sequence-dependent, or protein-induced bending of DNA, a large amount of work has been carried out to characterize it on structural and thermodynamical grounds (15,16). It is now commonly accepted that the bendability of specific DNA sequences relates to their capacity to be bent under the action of DNA-binding proteins. The bendability of these sequences may stem from their intrinsic bend, their low bending modulus or some specific breathing behavior of the DNA duplex structure that would facilitate its interactions with proteins. A DNA-analysis server based on the bending propensities of tri-nucleotides, that were deduced from DNase I digestion data (17), can be used to predict DNA structure from sequence and get an estimation of the foreseen DNA bend angles (18). However, characterizing rigorously the local bending of DNA molecules is a crucial issue that remains highly challenging (13,19,20).

To probe the DNA bendability experimentally, the most popular but complex technique remains the DNA cyclization method which provides a measure of the efficiency of cyclization of DNA fragments in presence of DNA ligase (21,22). However, this method does not permit to distinguish between changes in bending modulus that could be due to either permanent or transient structural defects and the presence of a local bend, a question for which other experimental strategies are needed. Another indirect approach consists in carrying out gel shift electrophoresis experiments using DNA molecules with several intrinsic

\*To whom correspondence should be addressed. Tel: +33 5 61 17 54 68; Fax: +33 5 61 17 59 94; Email: tardin@ipbs.fr

<sup>†</sup>These authors contributed equally to the paper as first authors.

sically bent sequences in-phase and in opposition of phase or circularly permuted DNA fragments (23–25). Though easy to handle, this technique can only provide rough estimates of bending angle and may be difficult to employ in all cases. Both Nuclear Magnetic Resonance (NMR) and X-ray require expensive equipment and tedious sample preparation and analysis procedures. They are therefore not routinely employed to detect and quantify the angle of an unknown DNA local bend but to bring details at the base-pair (bp) scale of a DNA structure already known to exhibit a local bend (26,27). Bend angles have also been deduced from more indirect techniques based on distance-dependent processes such as Fluorescence Resonance Energy Transfer (FRET) (28,29) or Plasmon resonance coupling of nanoparticles (30). Note that for these four last methods, the investigations are restricted to DNA molecules much shorter than the persistence length. Bending angles can also be extracted from single-molecule force-extension curves at large forces, as first proposed in (31). In this approach, the bending angle is inferred from the apparent persistence length, itself being a parameter used to fit experimental force-extension curves. However, extracting accurate values of the persistence length in this context presents several inherent difficulties, notably the fact that the fitting equation is supposed to be valid in the large-force regime where non-linear stretching should also be taken into account (32). More recently direct visualizations by Atomic Force Microscopy (AFM) (33,34) and cryo-electron microscopy (35) have given quantitative measurements of the bend angles. However these methods have limitations. Though AFM apparently gives the most direct access to the bend angles, the measurements are potentially biased by the sample preparation (36,37). By contrast, the technically demanding cryo-EM is supposed to preserve close-to-native state of the DNA complex, but it may also induce biases due especially to the sample confinement into a 50-nm-thick layer (35).

In the present work, we propose a physical method for the measure of DNA bend angles in a single DNA molecule and DNA-protein complex, which leaves intact the DNA conformation by ensuring minimal interaction with surfaces or tagged particles. This method combines our recently developed High Throughput Tethered Particle Motion (HT-TPM) technique and analytical modeling. HT-TPM enables the tracking of the conformational dynamics of hundreds of single DNA molecules in parallel, free to fluctuate in solution (38) (Figure 1A, Supplementary Information video). To extract the bend angles from HT-TPM data, we developed a simple analytical formula based on a kinked Worm-Like Chain (WLC) model that we validated on simulated data. Applied to constructs made of 575 bp DNAs with in-phase assemblies of 1 to 7 6A-tracts, we find that the sequence CA<sub>6</sub>CGG has an intrinsic bend angle of  $19^\circ \pm 4^\circ$ . In addition, the slight difference between our experimental data and our analytical model for a regular DNA suggests that even a DNA molecule with a randomly chosen sequence may contain a global curvature. We further apply our procedure to TPM measurements from the literature and demonstrate that the angles of bends induced by proteins, such as Integration Host Factor (IHF) can be reliably evaluated as well.

## MATERIALS AND METHODS

### DNA constructs

DNA molecules were produced by PCR (oligonucleotides (Sigma-Aldrich): Biot-F575 ATAAGGGCGACACG-GAAATG and Dig-R575 CGTGCAAGATTCCGAAT-ACC) on pTOC plasmids, derived from pBR322 (39). Synthetic (GeneScript) DNA molecules (88 bp) containing increasing number of A6-tracts (0 to 7) were synthesized (Figure 1B) and were inserted between the HindIII and SalI restriction sites of pBR322. PCR products were purified as described in Diagne *et al.* (40).

### HT-TPM setup and procedure

HT-TPM on chip assembly and the experimental setup are such as described in (38). The DNA/particle complexes were visualized in TPM buffer with an ionic strength of 165 mM (1.06 mM KH<sub>2</sub>PO<sub>4</sub>, 3.00 mM Na<sub>2</sub>HPO<sub>4</sub>, 154 mM NaCl, 1 mg.ml<sup>-1</sup> Pluronic F127, 0.1 mg.ml<sup>-1</sup> BSA) at  $21 \pm 1^\circ\text{C}$  using a dark-field microscope (Axiovert 200, Zeiss). Acquisitions of 5 min were performed at a recording rate of 25 Hz with acquisition time of 40 ms on a CMOS camera Dalsa Falcon 1.4M100. The field of observation covers an area of  $\sim 215 \times 160 \mu\text{m}$ . Details can be found in Supplementary Information text.

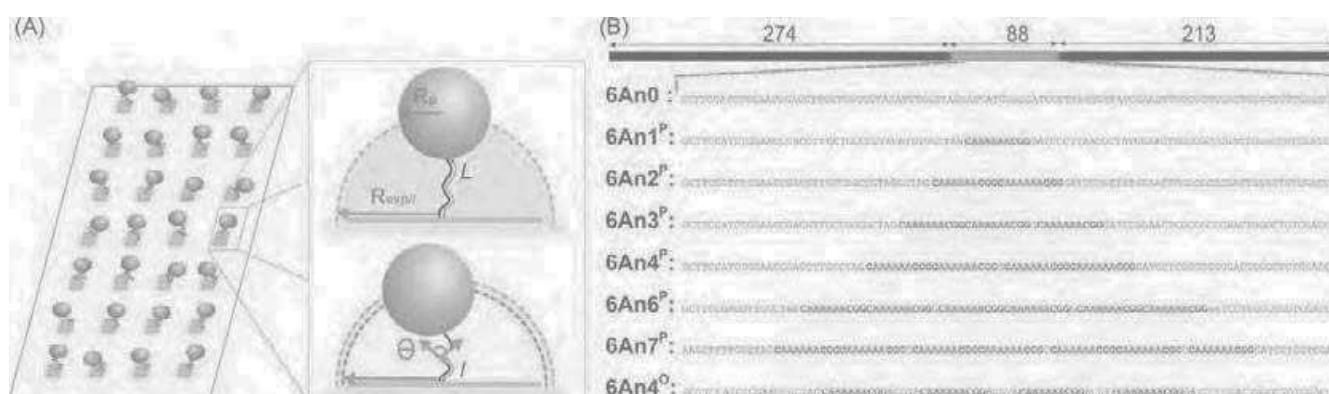
The software Nanomultiplex co-developed with Magellium Toulouse (request should be addressed at info@magellium.fr) tracks in real time the positions of all the particles using the centroid method, averages these absolute positions on a 5 s window giving access to the anchoring point of the DNA molecule, calculates the 2D-vector positions of the bead  $\vec{R}_{\text{exp||raw}}$  relative to the anchoring point of the DNA which corrects for experimental drift, calculates the asymmetry factor of the particle trajectories (41), and the amplitude of motion of the particle defined as  $\sqrt{\langle \vec{R}_{\text{exp||raw}}^2 \rangle}$  and noted  $R_{\text{exp||raw}}$ . In a

general manner, we will use indifferently  $\sqrt{\langle \vec{R}^2 \rangle}$  and  $R$  in the following. The averages, performed in the calculation of the asymmetry factors and amplitudes of motion, are taken over a sliding window of 5 s along the time trace. We invite the reader to refer to (42) for the detailed calculations of  $R_{\text{exp||raw}}$ .

Finite exposure time of detectors,  $T_{\text{ex}}$ , equal here to 40 ms, can lead to a blurring effect in single molecule (or particle) tracking experiments, as investigated for example in (42). The correlation time of the positions  $\tau_{\text{||}}$ , of about 20 ms, was calculated for each DNA-particle complex and then injected in the following equation:

$$R_{\text{exp||}} = R_{\text{exp||raw}} \left[ 2 \frac{\tau_{\text{||}}}{T_{\text{ex}}} - 2 \left( \frac{\tau_{\text{||}}}{T_{\text{ex}}} \right)^2 \left( 1 - e^{-\frac{T_{\text{ex}}}{\tau_{\text{||}}}} \right) \right]^{-1/2} \quad (1)$$

to correct the amplitudes of motion recorded on each trace from blurring. In order to quantify the small differences expected on  $R_{\text{exp||}}$ , we set up a two-step procedure that is described in detail in the Supplementary Information text.



**Figure 1.** (A) Schematic HT-TPM setup to measure the apparent length of a DNA molecule. The amplitude of motion,  $R_{\text{exp||}}$ , of a particle tethered to a surface by a DNA with a contour length  $L$  depends on its effective length which varies with the angle  $\theta$  of a bend located at distance  $l$  from one end chosen nearby the center of the molecule. For high-throughput measurements, individual DNA-particle complexes are immobilized on an array of functionalized sites. See ‘Materials and Methods’ section for details. (B) Representation of a typical 575-bp-long DNA fragment used in this study with the variable 88 bp central region (sizes are indicated in bp, details in Supplementary Information). The central region contains  $n$  copies of a  $\text{CA}_6\text{CGG}$  sequences (in red), which are predicted to be bent (see text). The sequences of the 88-bp-long insert of the eight DNA molecules studied here are represented with the  $\text{CA}_6\text{CGG}$  sequences. Each sequence is named as a function of the number of A contained in the  $\text{CA}_6\text{CGG}$  (6A), followed by the number of repeats (nX) and the phasing (in exponent, P: in-phase, O: in opposition of phase).

### TPM simulations

We performed Kinetic Monte Carlo simulations on the particle-DNA complex to predict the particle to anchor 2D-distance. We invite the reader to refer to (42) for the details of these Kinetic Monte Carlo simulations.

The bent sequences used in the experiments are simulated by setting a fixed angle between three successive monomers located at the center of the DNA molecule. A full range of angles were studied in successive simulations: 0, 18, 30, 45, 50, 60, 72, 90, 120 and 180°. The 2D-vector of the particle position  $\vec{R}_{\text{sim||}}$  is measured throughout simulations and utilized to estimate the amplitude of motion defined as  $\sqrt{\langle \vec{R}_{\text{sim||}}^2 \rangle}$ , the average being taken along the trajectory.

### Influence of the particle size on the angle measurements

The particle size cannot be known exactly. Indeed, on the one hand, there is a limited control of their radius by the manufacturer leading to an uncertainty of 3 nm. On the other hand, the subsequent functionalization of the particles confer them a slightly larger radius  $R_p$  (Figure 1A). Indeed, the layer of antibody used to connect the DNA molecule to the particle is expected to be a few nanometers wide. Since in Equation (2),  $R_p$  precisely refers to the distance between the particle center and the extremity of the DNA molecule, and being not able to exactly infer it, we chose to explore two cases:  $R_p = 150$  nm, previously used, and  $R_p = 155$  nm, the real values probably dwelling between both values.

For  $R_p = 155$  nm,  $\theta_l$  was found to be equal to  $17^\circ \pm 2^\circ$ , while for  $R_p = 150$  nm,  $\theta_l = 15^\circ \pm 2^\circ$ . The bend angle values obtained considering or not a 5 nm increase of the effective radius of the particles in this precise case cannot be distinguished. Though particles with a well-defined size would lead to a more precise bend angle, the uncertainty on the exact value of  $R_p$  does not appear to be a critical issue in a typical case such as the one described here.

## RESULTS

### A new experimental strategy

The presence of a localized bent structure within a DNA molecule is expected to induce the reduction of the apparent end-to-end distance of the entire DNA molecule. Such an effect will be much easier to detect on short DNA molecules. Such an effect will be much easier to detect on short DNA molecules. To measure it, we chose 575-bp-long DNA molecules that were immobilized on a functionalized coverslip by one end, and attached to a 150 nm radius particle at their other end permitting their video tracking (Figure 1A). This experimental design represents a good compromise for an easy particle detection and a nearly force-free measurement, as the effective force exerted by the particle on the DNA is only a fraction of pN (43). In that way, the 2D projection of the particle displacement relative to the anchoring point of the DNA molecule gives access to its root-mean-squared end-to-end distance projected on the grafting surface, noted  $R_{\text{exp||raw}}$ , which depends on the length and the conformational state of the monitored molecule.  $R_{\text{exp||raw}}$  was corrected for the blurring effect, caused by the acquisition system, to obtain  $R_{\text{exp||}}$ , using Equation (1) of ‘Materials and Methods’ section. TPM is capable of revealing changes in tether length equivalent to an apparent contour length as small as 100 bp (42,44). To obtain a good precision, inferior to 1 nm, a large amount of experimental data (see Table 1) is required because of the intrinsic dispersion of data due to an unavoidable variability of the DNA-particle and DNA-substrate links. To do so, we take advantage of our recently developed biochip that permits us to accumulate acquisitions on several hundreds of single DNA molecules in parallel by HT-TPM with a typical error on  $R_{\text{exp||}}$  equal to 0.4 nm (38).

In order to evaluate the capacity of this technique to detect and quantify local bending angles, we produced a series of DNA molecules based on a unique plasmid series. It incorporates a central 88 bp region, smaller than the DNA

**Table 1.** Amplitude of motion,  $R_{\text{exp||}}$ , corrected from the blurring effect and end-to-end distance,  $R_{\text{DNA}}$ , extracted with a particle radius  $R_p = 150$  nm for the set of constructs

DNA samples	Number of trajectories	$R_{\text{exp  }}$ (nm)	Uncertainty (nm)	$R_{\text{DNA}}$ (nm)
6An0	3496	149.7	0.2	105.4
6An1 <sup>P</sup>	2728	149.1	0.2	104.1
6An2 <sup>P</sup>	2354	149.7	0.2	105.4
6An3 <sup>P</sup>	2904	149.6	0.2	105.2
6An4 <sup>P</sup>	2604	146.5	0.2	98.5
6An6 <sup>P</sup>	348	145.2	0.4	95.5
6An7 <sup>P</sup>	431	140.4	0.4	84.1
6An4 <sup>O</sup>	2990	151.2	0.2	108.6

persistence length ( $L_p \approx 150$  bp), containing one to seven CA<sub>6</sub>CGG sequences in-phase or in opposition of phase (Figure 1B and Materials and Methods section). We chose the CA<sub>6</sub>CGG sequence, known to be a sequence inducing a large bend (45). The in-phase A-tracts are located every integer number of helix turns, whereas those in opposition of phase are located every half-integer number of helix turns (Figure 1B).

By HT-TPM, we measured  $R_{\text{exp||raw}}$  and calculated  $R_{\text{exp||}}$ , the amplitude of motion, for this series of 575-bp-long DNA molecule (Table 1).

#### A kinked Worm-Like Chain model to analyze TPM data

Obtaining the bend angle from  $R_{\text{exp||}}$  requires the calculation of the root-mean-square end-to-end distance of the DNA molecule,  $R_{\text{DNA}}$ , and an appropriate theoretical model giving the variation of  $R_{\text{DNA}}$  as a function of a local bend angle, denoted by  $\theta$ . To validate this analytical tool, we simulated a DNA-particle complex in a TPM setup (see Materials and Methods section and Supplementary Information text) where a bend of fixed angle  $\theta$  varying from 0 to 180° was incorporated in the middle of the DNA molecule. The corresponding projection of the mean-square end-to-end distance  $R_{\text{sim||}}$  was computed accordingly, as sketched in Figure 1A.

$R_{\text{DNA}}$  was extracted from these numerical results by correcting for the effects of the particle and of the glass substrate (42,43). To achieve it, we explored two strategies. The first one is a minimal model in which the particle and the DNA molecule are considered statistically independent and the effect of the substrate is ignored. This last assumption is correct when the particle is very small, ( $R_p \ll R_{\text{DNA}}$ ), and the DNA molecule is either very long ( $L \gg L_p$ ) or very short ( $L \ll L_p$ ), leading to  $\langle R_{\text{||}}^2 \rangle = \frac{2}{3} [R_{\text{DNA}}^2 + R_p^2]$  and thus to:

$$R_{\text{DNA}} = \left[ \frac{3}{2} \langle R_{\text{||}}^2 \rangle - R_p^2 \right]^{1/2} \quad (2)$$

However, it has been shown that the effect of the particle can matter, because the hard-core interaction between the particle and the substrate reduces the number of degrees of freedom accessible to the molecule. The second strategy relies on a more sophisticated protocol proposed by Segall *et al.* in (43) to correct for the effects of the particle and of the glass substrate for very long DNA. We refer to this approach as Segall's method. It consists in solving the equation:

tion:

$$\frac{3}{2} \frac{\langle R_{\text{||}}^2 \rangle}{\langle R_{\text{DNA}}^2 \rangle} = 1 + \frac{2N_R}{\sqrt{\pi} \operatorname{erf}(N_R)} \quad (3)$$

where  $\operatorname{erf}$  is the error function and  $N_R = \sqrt{6} R_p / \sqrt{\langle R_{\text{DNA}}^2 \rangle}$ . To use this Equation (3) in the present context, we crudely extrapolated the results of (43), obtained for the Gaussian chain case ( $L \gg L_p$ ), to the semi-flexible regime ( $L \sim L_p$ ), by replacing  $2LL_p$  by  $\langle R_{\text{DNA}}^2 \rangle$ . Note that Equation (2) is nothing but the expansion at order 1 in  $N_R$  of Equation (3).

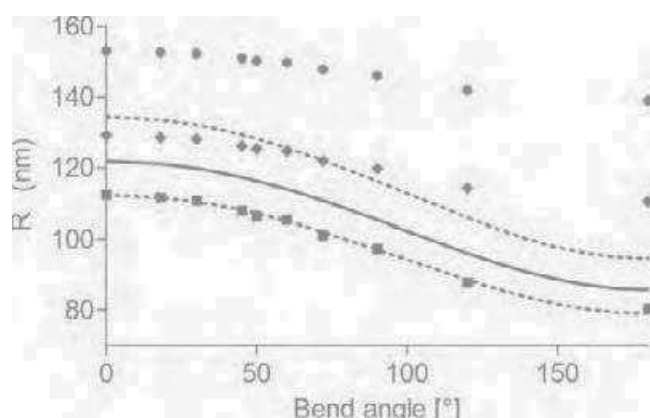
To obtain the bending angle, we fit the corrected data using a WLC model (46) on a polymer of length  $L$  with persistence length  $L_p$ , with a bend located at distance  $l$  from one end which locally induces a curvature with an angle  $\theta$  (see Figure 1A). The mean-squared end-to-end distance is given by (Supplementary Information text, note that a similar formula has been obtained in 2 dimensions in (33)):

$$\begin{aligned} \langle \vec{R}^2 \rangle(\theta) = & 2L_p^2 \left[ \frac{L}{L_p} - 2 + e^{-\frac{l}{L_p}} + e^{-\frac{L-l}{L_p}} \right] \\ & + \cos(\theta) \left[ 2L_p^2 \left( 1 - e^{-\frac{L-l}{L_p}} - e^{-\frac{l}{L_p}} + e^{-\frac{L}{L_p}} \right) \right] \end{aligned} \quad (4)$$

For  $L = 575$  bp and  $l = L/2$ , one derives from Equation (4) the fitting formula,  $R_{\text{WLC}}(\theta) = \sqrt{\langle \vec{R}^2 \rangle} = D\sqrt{1 + 0.342\cos\theta}$ . In Figure 2, both methods were applied to the numerical data. For example, for  $\theta = 0$ , the true WLC value of  $R_{\text{DNA}}$  is  $R_{\text{WLC}}(\theta = 0) = 121.9$  nm (see Materials and Methods section), in-between the values obtained with Segall's model and the minimal one. Whereas the minimal method underestimates  $R_{\text{WLC}}$ , Segall's one overestimates it, both methods leading to comparable relative errors of about 10%. We thus choose  $D$  as a free fitting parameter in order to account for an offset at  $\theta = 0$ . One observes in Figure 2 that whereas the minimal method gives a very satisfying fit, Segall's one leads to a worse one. Hence extrapolating Segall's calculation to the case of semi-flexible DNA molecules including a local bend appears to be less adapted than the minimal model. We thus use the minimal model in the remainder of this paper.

#### Bending angle measurements on a regular DNA

For the set of DNA molecules containing series of in-phase 6A-tracts (Table 1), we found by HT-TPM that  $R_{\text{exp||}}$  diminishes from 149.7 to 140.4 nm. When the initial 88 bp internal fragment of the DNA molecules (6An0) was replaced



**Figure 2.** Simulated HT-TPM amplitudes of motion  $R_{\text{sim}}$  (●), together with the corrected values  $R_{\text{DNA}}$  obtained through Segall's method (◆) and the minimal one (■); see text for details. The DNA molecule length is  $L = 575$  bp and the particle radius is  $R_p = 150$  nm, as in the experiments. Fits of the numerical data, corrected by both the Segall's method and the minimal one, are shown as dotted lines using Equation (4). The curve corresponding to the WLC model, with no adjusted parameters, is shown as a solid line.

by a fragment containing only one (6An1<sup>P</sup>), two (6An2<sup>P</sup>) or three 6A-tracts (6An3<sup>P</sup>) in-phase,  $R_{\text{exp}}$  remained in a range of 0.6 nm from its initial value (no 6A-tract), which appears to be very close to the 0.4 nm uncertainty range we estimated for our HT-TPM measurements according to the method described in the Materials and Methods section.  $R_{\text{exp}}$  decreased down to 146.5 nm for DNA molecules with four 6A-tracts in-phase (6An4<sup>P</sup>), 145.2 nm with six 6A-tracts in-phase (6An6<sup>P</sup>) and 140.4 nm with seven 6A-tracts in-phase (6An7<sup>P</sup>). These decreases in  $R_{\text{exp}}$  might stem from an intrinsic bend of the 6A-tract or a decrease in bending modulus of this sequence. In this latter case, one would expect the insertion of four 6A-tracts (6An4<sup>O</sup>) in opposition of phase to lead to a decrease in  $R_{\text{exp}}$  similar to the one measured on 6An4<sup>P</sup>. In fact, 6An4<sup>O</sup> showed no decrease in  $R_{\text{exp}}$  as it was found equal to 151.2 nm. Taken together, these results prove that, beyond four 6A-tracts,  $R_{\text{exp}}$  decreased when the number of 6A-tracts in-phase increases due to an intrinsic bend of the 6A-tract.

In order to calculate  $R_{\text{DNA}}$ , we injected these  $R_{\text{exp}}$  values in Equation (2). We assume then that each of the  $n$  successive 6A-tracts inserts in-phase imposes the same bending angle  $\theta_1$  and postulate as a first order simplification hypothesis that  $\theta = n\theta_1$ . With  $L = 575$  bp,  $l = 318$  bp, we now obtain the following equation:

$$R_{\text{WLC}}(n) = D[1 + 0.338\cos(n\theta_1)]^{1/2} \quad (5)$$

The fits of  $R_{\text{DNA}}$  shown in Figure 3 (black symbols and fits for  $R_p = 150$  nm) are reasonably good leading to  $D = 92 \pm 2$  nm and  $\theta_1 = 15^\circ \pm 2^\circ$  for the value of the bend angle of each CA<sub>6</sub>CGG sequence. In the Materials and Methods section, we studied the influence of small variations of the particle size on this result and found comparable results within error bar.

### Influence of a global DNA curvature on the angle measurements

The experimental values of Figure 3 suggested that  $R_{\text{DNA}}$  might be non-monotonous when  $n$  increases, contrary to what is expected from Equation (5). One reason for this observation could be the existence of an intrinsic curvature of the DNA molecule. To account for it, we have chosen to refine our analysis and use the more general fitting form:

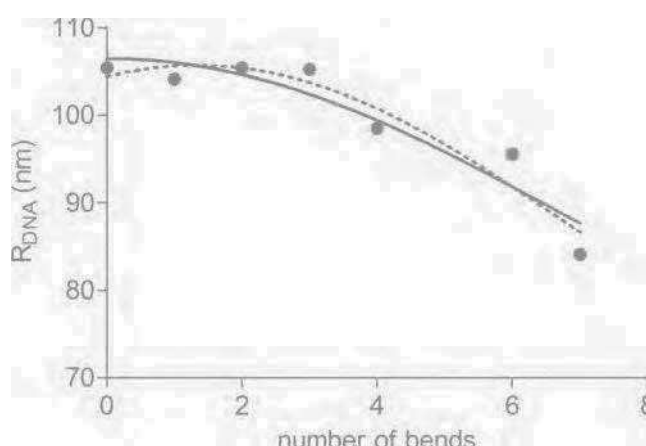
$$R_{\text{WLC}}(n) = D[1 + 0.338\cos(n\theta_1 - \theta_0)]^{1/2} \quad (6)$$

where  $\theta_0$  mimics the whole intrinsic bend of the  $n = 0$  molecule. For sake of simplicity, we implicitly assume that it is accumulated at the same position as the bends when  $n > 0$  (17,23). The resulting fits, displayed in Figure 3 (dotted lines), are slightly better than before.

To sum up, the proposed refinement suggests that the true bending angle  $\theta_1$  is  $19^\circ \pm 4^\circ$  per bend, while the global bend of the DNA molecule would have an angle  $\theta_0$  between  $25^\circ \pm 20^\circ$ . Although the fit including a global curvature of the DNA molecule seems better when looking at Figure 3, this new parameter  $\theta_0$  can only be determined with a poor precision. This is due to the fact that this angle is actually delocalized on the whole 575 bp sequence. A more sophisticated model is needed to take this point better into account.

### Influence of the bending modulus of the bent sequence on the angle measurements

Alternatively, it was in principle possible that the increase in the DNA end-to-end distance observed at small values of  $n$  was associated with an increase of the molecular rigidity of the inserts. We analytically derived the corrections to  $R_{\text{DNA}}$  due to an insert of finite length and of increased rigidity using the WLC model (Supplementary Information text). At small insert length, the increase is linear with the insert length (Supplementary Figure S1). At first order, this effect is additive with the effect of a bend. The corresponding fits



**Figure 3.** Experimental data after correction of the particle effects for the labeling particle  $R_p = 150$  nm (●) using the minimal model, and two series of fits with  $D$  and  $\theta_1$  (straight line), or  $D$ ,  $\theta_1$  and  $\theta_0$  (dotted line) as free parameters. Error bars are smaller than the symbol size. Using Equation (5), one finds:  $D = 92.0$  nm,  $\theta_1 = 15^\circ$  for  $R_p = 150$  nm. Using Equation (6), one finds  $D = 91$  nm,  $\theta_1 = 19^\circ$ ,  $\theta_0 = 25^\circ$  (see Table S3 for details).



are displayed in Supplementary Figure S2 and Table S3. The obtained values of  $\theta_1$  were close to the previous ones, thus showing that this alternative approach does not differ significantly from the previous one. This could be anticipated since a linear expansion of Equation (6) adds a linear term into Equation (5) if one assumes that  $\theta_0 \ll 1$ . As a result, the combined HT-TPM/kinked-WLC approach might not easily discriminate between an increase of rigidity of the inserts and an intrinsic bend. However, when fitting the data of the DNA molecules, containing 0 to 4 6A-tracts inserts, with this additional linear term, we found a linear increase of the amplitude of motion of 1.9 nm per 10 bp 6A-tract. It would correspond to a persistence length of about 89 nm for this bent sequence. There exists another way to evaluate the rigidity contribution of the 6A-tracts. It consists in considering the data obtained for 6An0 (no 6A-tracts) and 6An4<sup>O</sup> (four 6A-tracts in opposition of phase). An increase in  $R_{\text{DNA}}$  of 3.2 nm (105.4 versus 108.6 nm; Table 1) was then measured and gave access to a persistence length equal to 129 nm for this 55 bp sequence by using the formula in Supplementary Information text. These persistence lengths are estimated with a low precision due to the poor sensitivity of  $R_{\text{DNA}}$  to an increase of  $L_p$  above 150 bp (Supplementary Figure S4). In any case, whatever the way used to estimate the persistence lengths of these 6A-tract sequences, values were very high compared to the 50 nm measured for random double strand DNA in the present salt conditions (47). As a result, an increase in rigidity of the 6A-tracts could not account for our data, and introducing a global bend  $\theta_0$  is probably more adapted to this case.

## DISCUSSION

### DNA bending can be measured using HT-TPM

In this paper, we develop a procedure to measure a bend angle localized inside long DNA molecules with a good accuracy. For that, we built a platform plasmid based on pBR322 that enabled us to generate several 575-bp-long DNA molecules with various 88-bp-long DNA constructs close to their center. We used 6A-tracts as a model of bent sequences and showed that, by assembling them every 10.5 bp, we could amplify the observed decrease in  $R_{\text{exp}}$  as a function of the number of bent sequences. Furthermore, making use of the HT-TPM, we are able to reliably detect variations of  $R_{\text{exp}}$  down to 2%. Using Equations (1) and (2), we corrected experimental raw data for the effects of both camera averaging and particle radius and obtain  $R_{\text{DNA}}$ , which was correctly fitted using a kinked WLC model (Equation (4)). We thus got access to the bend angle of the chosen sequence  $\theta_1 = 19^\circ \pm 4$ .

To sum up, this entire procedure can be readily applied to quantify the bend related to a specific DNA structure together with an evaluation of the intrinsic curvature of the entire DNA molecule. To do so, our platform plasmid can be used to build an *ad hoc* series of DNA molecules with increasing numbers of the DNA sequences under study assembled in-phase. HT-TPM measurements have first to be corrected for the camera blurring effect using Equation (1) before being injected in Equation (2) to get the DNA end-to-end distance. The resulting series of data, DNA end-to-end distance as a function of the number of assembled se-

quences, are adjusted with Equation (4) where an intrinsic bend angle is incorporated if necessary.

### Comparison with other methods for the measurements of CA<sub>6</sub>CGG bend angles

Crothers *et al.* combined numerical simulations and cyclization experiments on DNA fragments of 105 and 210 bp lengths to get the relative bend induced by this sequence (48). They found a bend angle equal to  $19^\circ \pm 2^\circ$ . It should be stressed that cyclization is an indirect method which (i) enforces the looping of short DNAs ( $\approx L_p$ ) that can also be due to kinks or small denaturation bubbles (22), (ii) requires numerical simulations of the specific sequence with a large number of unknown parameters for the fitting procedure. In our case, we need at most three free parameters, namely  $D$ ,  $\theta_0$  and  $\theta_1$ .

More precise results were obtained by MacDonald *et al.* who used NMR spectroscopy on partially aligned DNA molecules (26). In particular, they found that the overall helix axis of the DNA dodecamer GGCA<sub>6</sub>CGG exhibits a bend of  $19^\circ \pm 1^\circ$  toward the minor groove of the A-tract. In addition, NMR results confirmed the major role played by the joint located between the 6A-tract and the next CG sequence that was found to be responsible for  $14^\circ$  of the  $19^\circ$  of the global bend angle of the dodecamer. NMR is limited to small DNA fragments and cannot be easily extended to DNA-protein complexes.

X-ray crystallography showed A-tracts structures with various conformations due to the crystal constraints (49). As it is complicated to determine the most representative one, we did not try to compare X-ray crystallography results to ours.

Experiments based on FRET also gave access to the bend angle value of the CA<sub>6</sub>CGG-CGA<sub>6</sub>CGG-CA<sub>6</sub>CGG sequence prepared in solutions with a broad range of salt concentration (28). The bend angle was deduced from the average dye-to-dye distance between two fluorophores located at the ends of the 31 bp DNA fragment. It was shown to vary from  $23^\circ \pm 4^\circ$  to  $41^\circ \pm 4^\circ$  with NaCl concentration increasing from 10 to 500 mM. At the ionic strength of 165 mM we carried out the experiments, Tóth *et al.* found a bend angle ranging between  $36^\circ \pm 4^\circ$  and  $41^\circ \pm 4^\circ$  which is in slight disagreement with the  $3 \times (19^\circ \pm 4^\circ) = 57^\circ \pm 7^\circ$  that we obtain. This discrepancy could be due to the sequence located on the middle which is slightly different with the flanking C being replaced by G. According to Koo *et al.* (45), the middle sequence should induce a bend smaller, up to 10%, than the two other ones. As a result, the total bend angle of the Toth's sequences is necessarily smaller than the one obtained by a 3-fold assembly of our sequence. Another source of discrepancy is that distance measurements by FRET are highly sensitive to the precise geometry of the attachment of the dyes at the extremity of the DNA molecules which is difficult to determine accurately. Moreover, the dyes are very close to the studied sequence and may interfere with it.

Finally, direct imaging of two to eight 6A-tracts assembled in-phase has been made by AFM (33). By measuring the planar end-to-end distance of these molecules, it was deduced a mean bend angle for each 6A-tract of about  $13.5^\circ$

with a precision of about  $1^\circ$ . The high precision that was obtained here does not take into account the possible bias induced by the mechanical constraints exerted onto the DNA molecule deposited on mica and visualized in air. Indeed the divalent ions that stick the DNA to the surface are likely to modify significantly the electrostatic interactions and therefore the DNA elastic properties (37).

Using HT-TPM to characterize the 6A-tracts enabled us to obtain values in reasonable agreement with those measured in these preceding studies by other methods. Besides, our measurements were carried out on DNA molecules containing the bent inserts far from the surface and the labeling particle, which thus do not interfere with the insert under examination. This is not possible with the previously mentioned techniques. Though very informative, most of them remain limited to the study of short DNA fragments. Furthermore, our underlying model is simple, appealing to elementary and robust polymer physics arguments, whereas interpreting data incoming from NMR, X-ray scattering, FRET or cyclization experiments is far more indirect and tedious. Finally, parallelizing single molecule experiments is a pre-requisite here, the very small variations of the amplitude of motion that we intend to monitor requiring intensive sampling.

### DNA heterogeneity and mechanical characterization

As expected the end-to-end distance of 6An0, calculated as  $R_{\text{DNA}}(\theta = 0)$ , is found notably smaller,  $107.7 \pm 7.1$  nm, than the expected WLC value, namely 121.9 nm. It suggests that the WLC model is not perfectly adequate to the DNA sample studied here. More importantly, whatever the equation we used to fit our HT-TPM data, the fits do not match exactly the experimental data within error bars. This can be due to an under-estimation of our confidence intervals. Even though increasing our sample sizes, we did not manage to reduce this discrepancy. We propose that it could be due to the intrinsic bend accumulated along the molecule, the additional effect of which would reduce the overall apparent end-to-end distance. Notably, when deleting some random DNA regions and replacing them by bent inserts, we conjecture that the deleted regions already bore some small but non-zero curvature, the effect of which can be either compensated or accentuated when replaced by an insert. This leads to small corrections of the end-to-end distance around the average trend described by Equation (4). This work thus highlights that a given sequence generically involves a small but non-vanishing intrinsic bend to DNA molecules, the amplitude of which typically grows like the square root of the molecule length, owing to the central-limit theorem, because the small bending angles related to individual base pairs add up independently along the chain. From a statistical physics perspective, our results point to the need to adapt statistical models to incorporate the fact that typical DNA molecules with a 'random' sequence adopt a curved shape rather than a straight one in the rigid regime  $L < L_p$  (13).

### Quantification of protein-induced DNA bends by HT-TPM

TPM is one of the very few methods permitting the monitoring of DNA looping. In addition, relying on the theo-

retical framework proposed here, we claim that HT-TPM can also reveal the protein-induced bending of the DNA interaction site and give access to the protein-induced bend angle.

When studying the interaction of a protein with a DNA molecule, a decrease in  $R_{\text{DNA}}$  may be observed after the addition of the protein on DNA molecules containing only one site of interaction (40,44,50–52). In this case, we can still evaluate the bend angle  $\theta$  from the two DNA end-to-end distance values obtained before and after addition of the protein by using Equation (7), derived from Equation (4) assuming that the global intrinsic bend  $\theta_0$  is negligible,

$$\cos(\theta) = 1 + \frac{e^{-\frac{L}{L_p}} + \frac{L}{L_p} - 1}{1 - e^{-\frac{L}{L_p}} - e^{-\frac{L}{L_p}} + \frac{L}{L_p}} \left( \frac{R_{\text{DNA}}^2(\theta) - R_{\text{DNA}}^2(0)}{R_{\text{DNA}}^2(0)} \right) \quad (7)$$

Note that if an intrinsic angle  $\theta_0$  was present before insertion of the sequence of interest, this formula gives access to the angle variation  $\theta - \theta_0$  instead of  $\theta$ .

As a test of its accuracy, we applied this simplified method to 6An4<sup>P</sup>, 6An6<sup>P</sup> and 6An7<sup>P</sup> and obtained a mean value of one 6A-tract insert equal to  $15^\circ \pm 2^\circ$ , the uncertainty is here calculated as the standard deviation of the data obtained for the three DNA molecules (Supplementary Table S5). This value is in good agreement with the value  $15^\circ \pm 2^\circ$  obtained from the fit performed on the entire set of data with  $\theta_1$  as the only free angle parameter ( $\theta_0 = 0$ ). In addition, using Equation (7) and considering equal to 2% the minimal variation in  $R_{\text{exp}}$  that can be detected with HT-TPM on this 575 bp DNA molecule, one can only measure bend angles larger than  $33^\circ$ . To measure smaller bend angles, the construction of a series of repeated sequence is required.

On the ground of this simple formula (Equation (7)), we analyzed the TPM results obtained by Mumm *et al.* (52) concerning the bend induced by the Integration Host Factor (IHF) that is known to bind to some specific binding sites and play thus a major role as architectural protein in prokaryotes. Upon binding of the protein to a single site of a  $L = 1943$ -bp-long DNA, the molecule behaved as a 1659-bp-long DNA which corresponds to  $R_{\text{DNA}}$  decreasing from  $R_{\text{DNA}} = 249$  to 229 nm. The binding site being located at  $l = 301$  bp away from one extremity, we evaluate the IHF induced-angle to be about  $180^\circ$  using Equation (7). This value is in very good agreement with the crystallography measurements that show that IHF provokes a bend with an angle of at least  $160^\circ$  (53), especially as it should be considered that the end-to-end distance tends to saturate for such high angles.

Our analytical models consider separately the cases of a bend with a fixed angle and of a local flexible hinge, while a great number of DNA-binding proteins are now considered to induce both effects. To account for these two mechanical changes occurring simultaneously as well as the extended deformation of the protein-DNA interaction site, a more precise theory is needed. Nevertheless, we believe that our method could still be applied to those cases and give access to some valuable though less precise information.

We can therefore conclude that the combined HT-TPM/kinked-WLC approach provides an efficient method

to estimate the angle of fixed local DNA bend either intrinsic to a sequence or induced by the binding of proteins. With many advantages over the existing methods, we believe that our approach will permit the refined characterization of DNA geometry in various contexts and to shed new light on DNA–protein complexes.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank François CORNET for his helpful discussions on the DNA construction. We belong to the GDR CellTiss.

## FUNDING

CNRS, University of Toulouse 3 and ANR-11-NANO-010 ‘TPM On a Chip’. Funding for open access charge: ANR-11-NANO-010 ‘TPM On a Chip’.

*Conflict of interest statement.* None declared.

## REFERENCES

- Borst, P. (1991) Why kinetoplast DNA networks? *Trends Genet.*, **7**, 139–141.
- Marini, J.C., Levene, S.D., Crothers, D.M. and Englund, P.T. (1982) Bent helical structure in kinetoplast DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **79**, 7664–7668.
- Wu, H.-M. and Crothers, D.M. (1984) The locus of sequence-directed and protein-induced DNA bending. *Nature*, **308**, 509–513.
- Horn, G.T. and Wells, R.D. (1981) The leftward promoter of bacteriophage lambda. Structure, biological activity, and influence by adjacent regions. *J. Biol. Chem.*, **256**, 2003–2009.
- Struhl, K. (1985) Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 8419–8423.
- Mizuno, T. (1987) Static bend of DNA helix at the activator recognition site of the ompF promoter in *Escherichia coli*. *Gene*, **54**, 57–64.
- Ross, W., Gosink, K.K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K. and Gourse, R.L. (1993) A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science*, **262**, 1407–1413.
- Gaal, T., Rao, L., Estrem, S.T., Yang, J., Wartell, R.M. and Gourse, R.L. (1994) Localization of the intrinsically bent DNA region upstream of the *E. coli* rrnB P1 promoter. *Nucleic Acids Res.*, **22**, 2344–2350.
- Tippner, D., Afflerbach, H., Bradacek, C. and Wagner, R. (1994) Evidence for a regulatory function of the histone-like *Escherichia coli* protein H-NS in ribosomal RNA synthesis. *Mol. Microbiol.*, **11**, 589–604.
- Iyer, V. and Struhl, K. (1995) Poly(dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. *EMBO J.*, **14**, 2570–2579.
- Hillebrand, A., Wurm, R., Menzel, A. and Wagner, R. (2005) The seven *E. coli* ribosomal RNA operon upstream regulatory regions differ in structure and transcription factor binding efficiencies. *Biol. Chem.*, **386**, 523–534.
- Segal, E. and Widom, J. (2009) Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.*, **19**, 65–71.
- Freeman, G.S., Lequieu, J.P., Hinckley, D.M., Whitmer, J.K. and de Pablo, J.J. (2014) DNA Shape Dominates Sequence Affinity in Nucleosome Formation. *Phys. Rev. Lett.*, **113**, 168101.
- Swinger, K.K. and Rice, P.A. (2004) IHF and HU: flexible architects of bent DNA. *Curr. Opin. Struct. Biol.*, **14**, 28–35.
- Maher, L.J. III (1998) Mechanisms of DNA bending. *Curr. Opin. Chem. Biol.*, **2**, 688–694.
- Privalov, P.L., Dragan, A.I. and Crane-Robinson, C. (2009) The cost of DNA bending. *Trends Biochem. Sci.*, **34**, 464–470.
- Bruckner, I., Sánchez, R., Suck, D. and Pongor, S. (1995) Sequence-dependent bending propensity of DNA as revealed by DNase I: parameters for trinucleotides. *EMBO J.*, **14**, 1812–1818.
- Vlahovicek, K., Kajan, L. and Pongor, S. (2003) DNA analysis servers: plot.it, bend.it, model.it and IS. *Nucleic Acids Res.*, **31**, 3686–3687.
- Gietl, A. and Grohmann, D. (2013) Modern biophysical approaches probe transcription-factor-induced DNA bending and looping. *Biochem. Soc. Trans.*, **41**, 368–373.
- Harteis, S. and Schneider, S. (2014) Making the bend: DNA tertiary structure and protein-DNA interactions. *Int. J. Mol. Sci.*, **15**, 12335–12363.
- Zhang, Y., Xi, Z., Hegde, R.S., Shakked, Z. and Crothers, D.M. (2004) Predicting indirect readout effects in protein–DNA interactions. *PNAS*, **101**, 8337–8341.
- Vafabakhsh, R. and Ha, T. (2012) Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cyclization. *Science*, **337**, 1097–1101.
- Zinkel, S.S. and Crothers, D.M. (1987) DNA bend direction by phase sensitive detection. *Nature*, **328**, 178–181.
- Lane, D., Prentki, P. and Chandler, M. (1992) Use of gel retardation to analyze protein–nucleic acid interactions. *Microbiol. Rev.*, **56**, 509–528.
- Wang, Q. and Calvo, J.M. (1993) Lrp, a major regulatory protein in *Escherichia coli*, bends DNA and can organize the assembly of a higher-order nucleoprotein structure. *EMBO J.*, **12**, 2495–2501.
- MacDonald, D., Herbert, K., Zhang, X., Pologruto, T., Lu, P. and Polgruto, T. (2001) Solution structure of an A-tract DNA bend. *J. Mol. Biol.*, **306**, 1081–1098.
- Nelson, H.C., Finch, J.T., Luisi, B.F. and Klug, A. (1987) The structure of an oligo(dA).oligo(dT) tract and its biological implications. *Nature*, **330**, 221–226.
- Tóth, K., Sauermann, V. and Langowski, J. (1998) DNA curvature in solution measured by fluorescence resonance energy transfer. *Biochemistry*, **37**, 8173–8179.
- Wozniak, A.K., Schröder, G.F., Grubmüller, H., Seidel, C.A.M. and Oesterhelt, F. (2008) Single-molecule FRET measures bends and kinks in DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 18337–18342.
- Reinhard, B.M., Sheikholeslami, S., Mastroianni, A., Alivisatos, A.P. and Liphardt, J. (2007) Use of plasmon coupling to reveal the dynamics of DNA bending and cleavage by single EcoRV restriction enzymes. *Proc. Natl. Acad. Sci.*, **104**, 2667–2672.
- Kulić, I.M., Mohrbach, H., Lobaskin, V., Thaokar, R. and Schiessel, H. (2005) Apparent persistence length renormalization of bent DNA. *Phys. Rev. E*, **72**, 041905.
- Manghi, M., Destainville, N. and Palmeri, J. (2012) Mesoscopic models for DNA stretching under force: New results and comparison with experiments. *Eur. Phys. J. E*, **35**, 110.
- Rivetti, C., Walker, C. and Bustamante, C. (1998) Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility. *J. Mol. Biol.*, **280**, 41–59.
- Buzio, R., Repetto, L., Giacomelli, F., Ravazzolo, R. and Valbusa, U. (2012) Label-free, atomic force microscopy-based mapping of DNA intrinsic curvature for the nanoscale comparative analysis of bent duplexes. *Nucleic Acids Res.*, **40**, e84.
- Bednar, J., Furrer, P., Katritch, V., Stasiak, A., Dubochet, J. and Stasiak, A. (1995) Determination of DNA persistence length by cryo-electron microscopy. Separation of the static and dynamic contributions to the apparent persistence length of DNA. *J. Mol. Biol.*, **254**, 579–594.
- Sorel, I., Piétrement, O., Hamon, L., Baconnais, S., Le Cam, E. and Pastré, D. (2006) The EcoRI–DNA complex as a model for investigating protein–DNA interactions by atomic force microscopy. *Biochemistry*, **45**, 14675–14682.
- Destainville, N., Manghi, M. and Palmeri, J. (2009) Microscopic mechanism for experimentally observed anomalous elasticity of DNA in two dimensions. *Biophys. J.*, **96**, 4464–4469.
- Plénat, T., Tardin, C., Rousseau, P. and Salome, L. (2012) High-throughput single-molecule analysis of DNA–protein interactions by tethered particle motion. *Nucleic Acids Res.*, **40**, e89.
- Watson, N. (1988) A new revision of the sequence of plasmid pBR322. *Gene*, **70**, 399–403.

40. Diagne,C.T., Salhi,M., Crozat,E., Salomé,L., Cornet,F., Rousseau,P. and Tardin,C. (2014) TPM analyses reveal that FtsK contributes both to the assembly and the activation of the XerCD-dif recombination synapse. *Nucl. Acids Res.*, **42**, 1721–1732.
41. Blumberg,S., Gajraj,A., Pennington,M.W. and Meiners,J.-C. (2005) Three-dimensional characterization of tethered microspheres by total internal reflection fluorescence microscopy. *Biophys. J.*, **89**, 1272–1281.
42. Manghi,M., Tardin,C., Baglio,J., Rousseau,P., Salomé,L. and Destainville,N. (2010) Probing DNA conformational changes with high temporal resolution by tethered particle motion. *Phys. Biol.*, **7**, 046003.
43. Segall,D.E., Nelson,P.C. and Phillips,R. (2006) Volume-exclusion effects in tethered-particle experiments: bead size matters. *Phys. Rev. Lett.*, **96**, 088306.
44. Pouget,N., Turlan,C., Destainville,N., Salomé,L. and Chandler,M. (2006) IS911 transpososome assembly as analysed by tethered particle motion. *Nucleic Acids Res.*, **34**, 4313–4323.
45. Koo,H.S., Wu,H.M. and Crothers,D.M. (1986) DNA bending at adenine. thymine tracts. *Nature*, **320**, 501–506.
46. Kratky,O. and Porod,G. (1949) Röntgenuntersuchung gelöster Fadenmoleküle. *Recl. Trav. Chim. Pays Bas*, **68**, 1106–1122.
47. Baumann,C.G., Smith,S.B., Bloomfield,V.A. and Bustamante,C. (1997) Ionic effects on the elasticity of single DNA molecules. *PNAS*, **94**, 6185–6190.
48. Koo,H.S., Drak,J., Rice,J.A. and Crothers,D.M. (1990) Determination of the extent of DNA bending by an adenine-thymine tract. *Biochemistry*, **29**, 4227–4234.
49. DiGabriele,A.D., Sanderson,M.R. and Steitz,T.A. (1989) Crystal lattice packing is important in determining the bend of a DNA dodecamer containing an adenine tract. *PNAS*, **86**, 1816–1820.
50. Tolić-Nørrelykke,S.F., Rasmussen,M.B., Pavone,F.S., Berg-Sørensen,K. and Oddershede,L.B. (2006) Stepwise Bending of DNA by a Single TATA-Box Binding Protein. *Biophys. J.*, **90**, 3694–3703.
51. Dixit,S., Singh-Zocchi,M., Hanne,J. and Zocchi,G. (2005) Mechanics of binding of a single integration-host-factor protein to DNA. *Phys. Rev. Lett.*, **94**, 118101.
52. Mumm,J.P., Landy,A. and Gelles,J. (2006) Viewing single  $\lambda$  site-specific recombination events from start to finish. *EMBO J.*, **25**, 4586–4595.
53. Rice,P.A., Yang,S., Mizuuchi,K. and Nash,H.A. (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell*, **87**, 1295–1306.

## SUPPORTING INFORMATION

### SI text

#### List of abbreviations

$\vec{R}_{exp  raw}$ :	2D-vector positions of the particle measured experimentally
$R_{exp  raw}$	Amplitude of motion of the particle defined as $\sqrt{\langle \vec{R}_{exp  raw}^2 \rangle}$
$R_{exp  }$	Amplitude of motion of the particle corrected from the blurring effect
$\vec{R}_{sim  }$	2D-vector positions of the simulated particle
$R_{sim  }$	Amplitude of motion of the simulated particle defined as $\sqrt{\langle \vec{R}_{sim  }^2 \rangle}$
$R_{DNA}$	End-to-end distance of the DNA molecule
$R_p$	Radius of the labelling particle
$L$	Contour length of the DNA molecule
$l$	Distance separating the kink to one extremity of the DNA molecule
$\theta$	Bend angle
$L_p$	Persistence length of the DNA molecule

#### DNA constructs

The sequences of the 88bp long insert (Fig. 1) are the following ones:

**6An0** : GCTTCCATCTCGAACCGACGTTGCTGGCCGTACATGTGGCTAGCACCATGGAGGATCCCTTAACGCTATGGAACTCGCCGCCGACTGGGCTGTCGACCG

**6An1<sup>P</sup>**: GCTTCCATCTCGAACCGACGTTGCTGGCCGTACATGTGGCTAGCAAAAAACGGGATCCCTTAACGCTATGGAACTCGCCGCCGACTGGGCTGTCGACCG

**6An2<sup>P</sup>**: GCTTCCATCTCGAACCGACGTTGCTGGCCGTAGGCTAGCAAAAAACGGCAAAAAACGGGATCCCGCTATGGAACTCGCCGCCGACTGGGCTGTCGACCG

**6An3<sup>P</sup>**: GCTTCCATCTCGAACCGACGTTGCTGGCGCTAGCAAAAAACGGCAAAAAACGGGCAAAAAACGGGATCCGGAACTCGCCGCCGACTGGGCTGTCGACCG

**6An4<sup>P</sup>**: GCTTCCATCTCGAACCGACGTTGGCTAGCAAAAAACGGCAAAAAACGGGCAAAAAACGGCAAAAAACGGGATCCTCGCCGCCGACTGGGCTGTCGACCG

**6An6<sup>P</sup>**: GCTTCCGACGTTGGCTAGCAAAAAACGGCAAAAAACGGGCAAAAAACGGGCAAAAAACGGGCAAAAAACGGGATCCTCGCCGCGTCGACCG

**6An7<sup>P</sup>**: AAGCTTTTGGCTAGCAAAAAACGGCAAAAAACGGGCAAAAAACGGGCAAAAAACGGGCAAAAAACGGGCAAAAAACGGGATCCTCGTCGAC

**6An4<sup>O</sup>**: GCTTCCATCTCGAACCGCTAGCAAAAAACGGACGTACAAAAACGGACGTACAAAAACGGACGTACAAAAACGGGATCCCGACTGGGCTGTCGACCG

## Assembly of the HT-TPM biochip

### *Patterning of anchoring sites on functionalized coverslips*

Regular arrays of rhodamine-labelled neutravidin (Molecular Probes) were obtained using a standard micro-contact printing protocol. Briefly, a stamp made of PDMS with squared pillars of 0.8  $\mu\text{m}$  size and 3  $\mu\text{m}$  pitch was inked with a 20  $\mu\text{g/mL}$  neutravidin solution in PBS (Euromedex) for 1 min, then washed with deionized water and dried under nitrogen flow. The stamp was then brought into close contact with epoxidized glass support for 1 min during which the protein is transferred onto the surface.

### *Formation of DNA-bead complexes*

Polystyrene carboxylated beads (Merck) of  $(150 \pm 3)$  nm radius were covalently coated with anti-digoxigenin antibodies (Roche) using EDAC (Sigma-Aldrich), activation and storage buffers (Ademtech). A 100 pM solution of suspended functionalized beads was mixed with an equal volume of a solution of 50 pM DNA bearing a digoxigenin on one end for 20 min at room temperature in PBS buffer supplemented with 1mg/mL of pluronic 127 (Sigma-Aldrich) and 0.1 mg/mL BSA (Sigma-Aldrich), noted TPM buffer. This lead to pre-assembled DNA-bead complexes.

### *Assembly of the fluidic observation chamber for HT-TPM experiments*

A 250  $\mu\text{m}$  thick silicone tape was cut and used as a spacer between the patterned coverslip and an epoxidized glass slide with 2 holes for inlet and outlet to obtain a working flow cell. The so-formed analysis chamber was rinsed and incubated with TPM buffer for 30 min at room temperature. The DNA-bead complexes solution was introduced in the flow cell and incubated over night at 4°C.

Prior to visualization, the flow cell was extensively rinsed by injecting 30 chamber volumes of TPM buffer. Then, for each condition, movies of 5 min were recorded on different zones in the same flow cell and analyzed. To ensure reproducibility, experiments were repeated on different days.

## HT-TPM experimental setup

### *Instrumentation for microscopy imaging*

The tethered beads were visualized using a dark-field microscope (Axiovert 200, Zeiss) equipped with a x32 objective and an additional x1.6 magnification lens and acquired for 5 min at room temperature, at a recording rate of 25 Hz and with a duration of acquisition of 40ms, on a CMOS camera Dalsa Falcon 1.4M100. The field of observation covers an area of  $\sim 215 \mu\text{m} \times 160 \mu\text{m}$ .

### *Single particle tracking*

The software Nanomultiplex co-developed with Magellium Toulouse (request should be addressed at [info@magellium.fr](mailto:info@magellium.fr)) tracks in real time the positions of all the particles using the centroid method, averages these absolute positions on a 5 s window giving access to the anchoring point of the DNA molecule, calculates the 2D-vector positions of the bead  $\vec{R}_{exp||raw}$  relative to the anchoring point of the DNA which corrects for experimental drift, calculates the asymmetry factor of the bead trajectories (S) (Blumberg S, Gajraj A, Pennington MW, Meiners J-C (2005) Three-Dimensional Characterization of Tethered Microspheres by Total Internal Reflection Fluorescence Microscopy. Biophys J 89(2):1272–1281), and the amplitude of motion of the particle defined as  $\sqrt{\langle \vec{R}_{exp||raw}^2 \rangle}$  and noted  $R_{exp||raw}$ . In a general manner, we will use indifferently  $\sqrt{\langle \vec{R}^2 \rangle}$  and  $R$  in the following.

The averages, performed in the calculation of the asymmetry factors and amplitudes of motion, are taken over a sliding window of 5 s along the time trace. We invite the reader to refer to (Plenat T, Tardin C, Rousseau P, Salome L (2012) High-throughput single-molecule analysis of DNA-protein interactions by tethered particle motion. Nucleic Acids Research 40(12):e89–e89) for the detailed calculations of  $R_{exp||raw}$  of the bead.

## Procedure of analysis for HT-TPM experimental data

In order to quantify the small differences expected on  $R_{exp||raw}$ , we set up a two-step procedure that is described in detail below. Briefly, it consists in selecting traces fulfilling several criteria of validity

and applying corrections for detector temporal averaging to their  $R_{exp//raw}$ . All this procedure was performed with homebuilt Mathematica scripts (available upon request).

### *Criteria of validity of the DNA-bead complexes*

First, we discard the trajectories that have mean asymmetry factors above 1.35, calculated as the average of the asymmetric factors measured along the time trace, or that have a mean amplitude of motion smaller than 1 nm or higher than 1000 nm. Then the probability distribution of the average of for each trajectory is built with the remaining trajectories and fitted by a Gaussian distribution centered on a mean value, called mean with a standard deviation (sd). As we noticed that a few trajectories had  $R_{exp//raw}$  averages standing out of the Gaussian distribution, we added a second step of validation to eliminate the misformed tethers with an average  $R_{exp//raw}$  outside the interval (mean  $\pm$  2.5 sd). Using this criterion, no more than 1.3% of valid trajectories were eliminated during this additional step.

In total, about 12% trajectories were eliminated and the final number of valid trajectories eventually ranged between 348 and 3496 (See Table 1), depending on the DNA construct.

### *Correction of time averaging effect*

Finite exposure time of detectors,  $T_{ex}$ , equal here to 40 ms, can lead to a blurring effect in single molecule (or particle) tracking experiments, as investigated for example in (Manghi M, et al. (2010) Probing DNA conformational changes with high temporal resolution by tethered particle motion. *Physical Biology* 7(4):046003). The correlation time of the positions, of about 20 ms, was calculated for each DNA-particle complex and then injected in Eq. 1 to correct the amplitudes of motion recorded on each trace.

$$R_{exp\parallel} = R_{exp\parallel raw} \left[ 2 \frac{\tau_{\parallel}}{T_{ex}} - 2 \left( \frac{\tau_{\parallel}}{T_{ex}} \right)^2 \left( 1 - e^{-\frac{T_{ex}}{\tau_{\parallel}}} \right) \right]^{-1/2} \quad (1)$$

As previously, the trajectories with an average  $R_{exp\parallel}$  outside the interval (mean  $\pm$  2.5 sd) were eliminated.



### *Calculation of the amplitude of the motion of an ensemble of particles*

Lastly, the experimental value of the amplitude of the motion of an ensemble of particles was obtained by fitting the probability distribution of  $R_{exp\parallel}$  with a Gaussian, which gave us its mean value as its center. The error on the amplitude of the motion of an ensemble of particles was calculated by using the bootstrap method of R software (R Foundation for Statistical Computing, Vienna, Austria). Doing so, we find a typical error of 0.4 nm (See Table 1).

## **TPM simulations**

### *DNA coarse-grained model*

We performed Kinetic Monte Carlo simulations (Newman MEJ, Barkema GT (1999) Monte Carlo Methods in Statistical Physics (Oxford University Press)) on the particle-DNA complex to predict the particle to anchor 2D-distance. The labeled DNA polymer is modeled as a chain of  $N$  connected small spheres of radius  $a$ , with a DNA contour length equal to  $L = 2a(N - 1)$ , and a larger particle of radius  $R_p = 150 \text{ nm} \geq a$  at its terminus. At this level of modeling, the internal structure of the double-stranded DNA is not considered and the persistence length,  $L_p$ , is averaged over the nucleotide sequence and taken equal to 147 bp. To model the angle imposed in the experiments by the  $n$  successive 10-bp CA<sub>6</sub>CGG inserts, we have incorporated a bend of fixed angle  $\theta$  between the three small beads located in the middle of the DNA molecule, which is simulated using a coarse-grained model. More precisely, the elastic term between triplets  $(i-1, i, i+1)$  of successive beads, making an angle  $\theta_i$ , namely  $U_i = \kappa (1 - \cos \theta_i)$ , is replaced by  $U'_i = \kappa (1 - \cos(\theta_i - \theta))$  for the central triplet where  $\kappa$  is the bending modulus and  $\kappa = 147k_B T$ . The polymer is grafted on a surface. Since the polymer motion is limited to the upper half-plane, we impose a “hard wall” boundary condition for monomer spheres and for the particle. All spheres interact via stretching and bending forces. We invite the reader to refer to (Manghi M, et al. (2010) Probing DNA conformational changes with high temporal resolution by tethered particle motion. Physical Biology 7(4):046003) for the details of these Kinetic Monte Carlo simulations.

### *Simulated bent DNA*

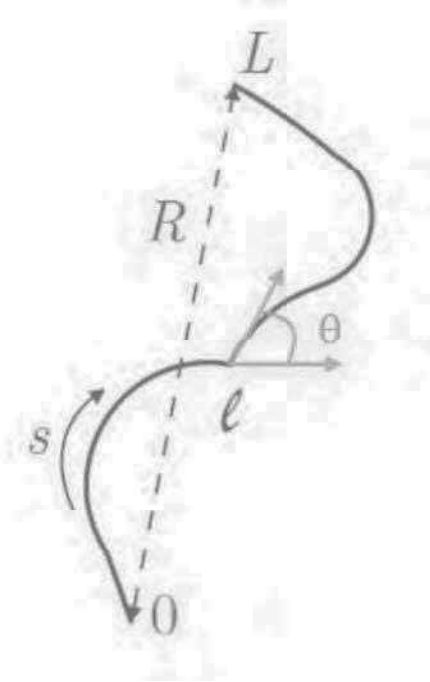
The bent sequences used in the experiments are simulated by setting a fixed angle between three successive monomers located at the center of the DNA molecule. A full range of angles were studied in successive simulations: 0, 18, 30, 45, 50, 60, 72, 90, 120 and 180 degrees.

### *Extraction/Calculation of the particle to anchor 2D-distance of the simulated particle*

The 2D-vector of the particle position  $\vec{R}_{sim\parallel}$  is measured throughout simulations and utilized to estimate the amplitude of motion defined as  $\sqrt{\langle \vec{R}_{sim\parallel}^2 \rangle}$ , the average being taken along the trajectory.

## Theory for a local bend

We consider a homogeneous polymer of length  $L$  using the Worm-Like Chain (WLC) model by Kratky and Porod with a persistence length  $L_p$  (Kratky O, Porod G (1949) Röntgenuntersuchung gelöster Fadenmoleküle. *Recl Trav Chim Pays-Bas* 68(12):1106–1122). We suppose that a kink is located at distance  $l$  from one end and locally induces a spontaneous curvature with an angle  $\theta$ , see below.



We denote the tangent vector at curvilinear position  $s$  by  $\overrightarrow{t(s)}$ , and the tangent-tangent correlation function is given by  $\langle \overrightarrow{t(s)} \cdot \overrightarrow{t(s')} \rangle = \exp[-|s - s'|/L_p]$ .

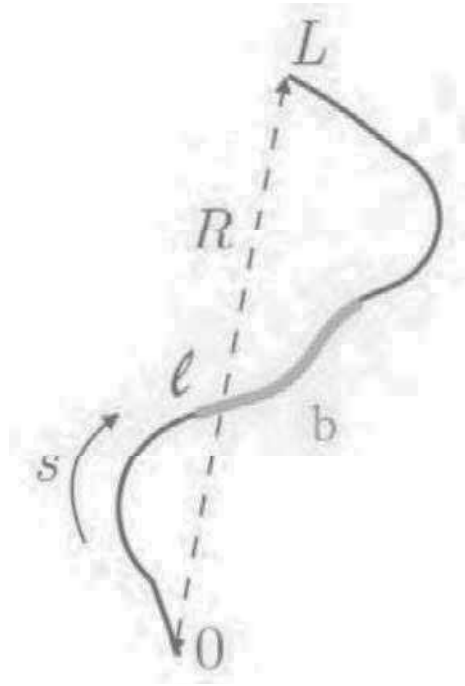
The mean-square end-to-end distance is defined as :  $\langle R^2 \rangle = \int_0^L ds \int_0^L ds' \langle \overrightarrow{t(s)} \cdot \overrightarrow{t(s')} \rangle$ .

Integration leads to the following result:

$$\langle R^2 \rangle = 2 L_p^2 \left[ \frac{L}{L_p} - 2 + e^{-\frac{l}{L_p}} + e^{-\frac{L-l}{L_p}} + \cos(\theta) \left( 1 - e^{-\frac{L-l}{L_p}} - e^{-\frac{l}{L_p}} + e^{-\frac{L}{L_p}} \right) \right]$$

## Theory for a local stiffer insert

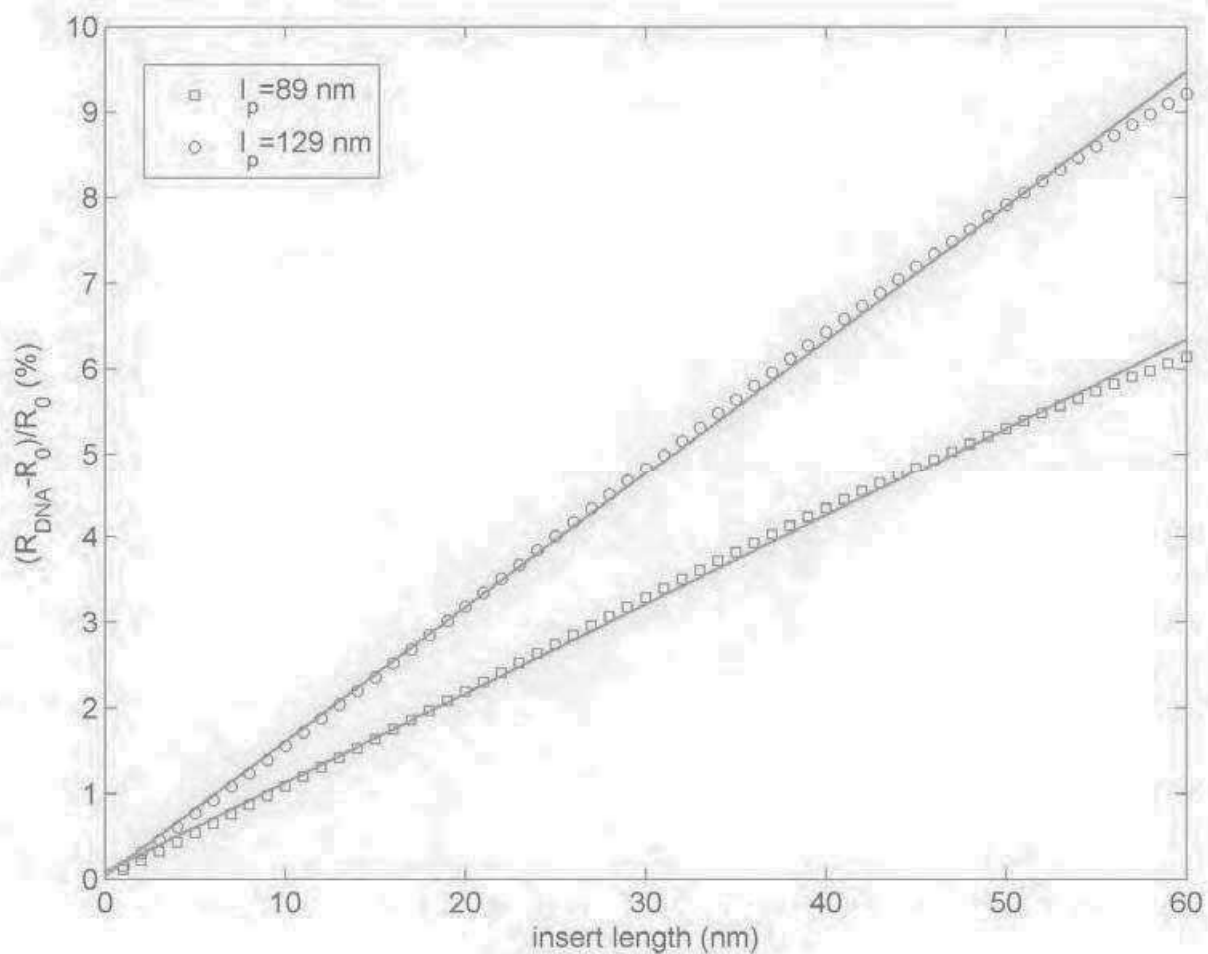
We use the same model and assume now that a stiffer insert of length  $b$  and persistence length  $L_{pb}$  is inserted at position  $l$ , see below:



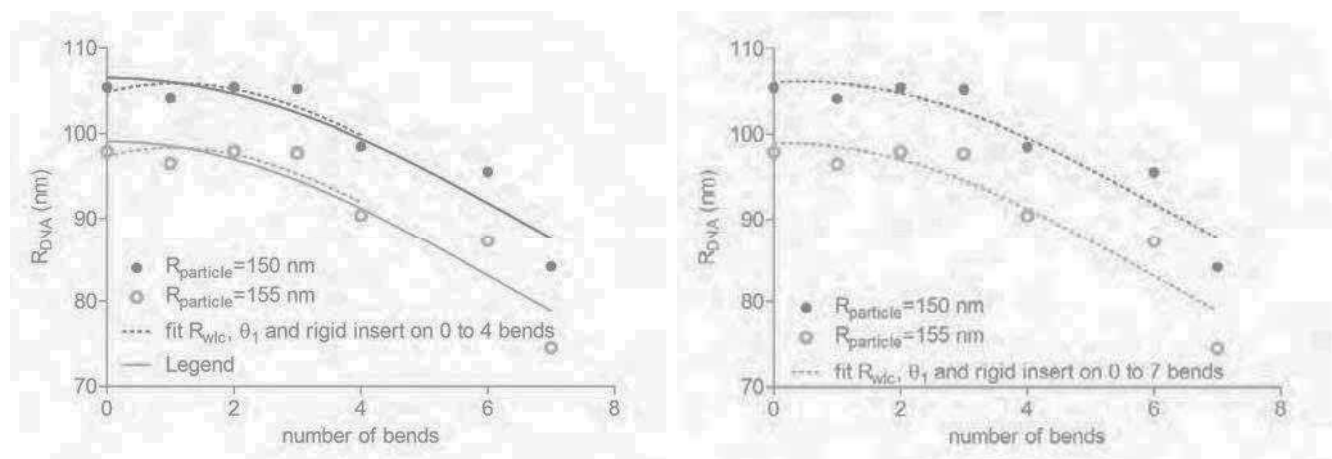
The mean-square end-to-end distance is again  $\langle R^2 \rangle = \int_0^L ds \int_0^L ds' \langle \vec{t}(s) \cdot \vec{t}(s') \rangle$ . To take into account the 3 parts of the DNA molecule, the calculation requires the equation to be cut into parts. After a straightforward integration, one gets:

$$\begin{aligned} \langle R^2 \rangle = & 2 L_P^2 \left[ \frac{L-b}{L_P} + e^{-\frac{l}{L_P}} + e^{-\frac{L-b-l}{L_P}} - 2 + e^{-\frac{b}{L_{pb}}} \left( 1 - e^{-\frac{l}{L_P}} \right) \left( 1 - e^{-\frac{L-b-l}{L_P}} \right) \right. \\ & \left. + 2 L_P L_{pb} \left( 1 - e^{-\frac{b}{L_{pb}}} \right) \left( 2 - e^{-\frac{l}{L_P}} - e^{-\frac{L-b-l}{L_P}} \right) \right] + 2 L_{pb}^2 \left( \frac{b}{L_{pb}} - 1 + e^{-\frac{b}{L_{pb}}} \right) \end{aligned}$$

**S1 figure:** linearity of  $R_{\text{DNA}}$  with the length of a rigid insert. Simulated data are represented by black markers and the linear fits by red straight lines.



**S2 figure:** complement to Fig. 3



Experimental data (symbols) obtained after deconvolution of the particle effects using the minimal model and considering the size of labelling particle equal to 150 nm (●) and 155 nm (○), and two series of fits with  $R_{WLC}$  and  $\theta_1$  (—) (already shown in figure 3), or  $R_{WLC}$ ,  $\theta_1$  and a rigidity term  $\alpha$  (...) as free parameters. The results are shown in S3 Table.

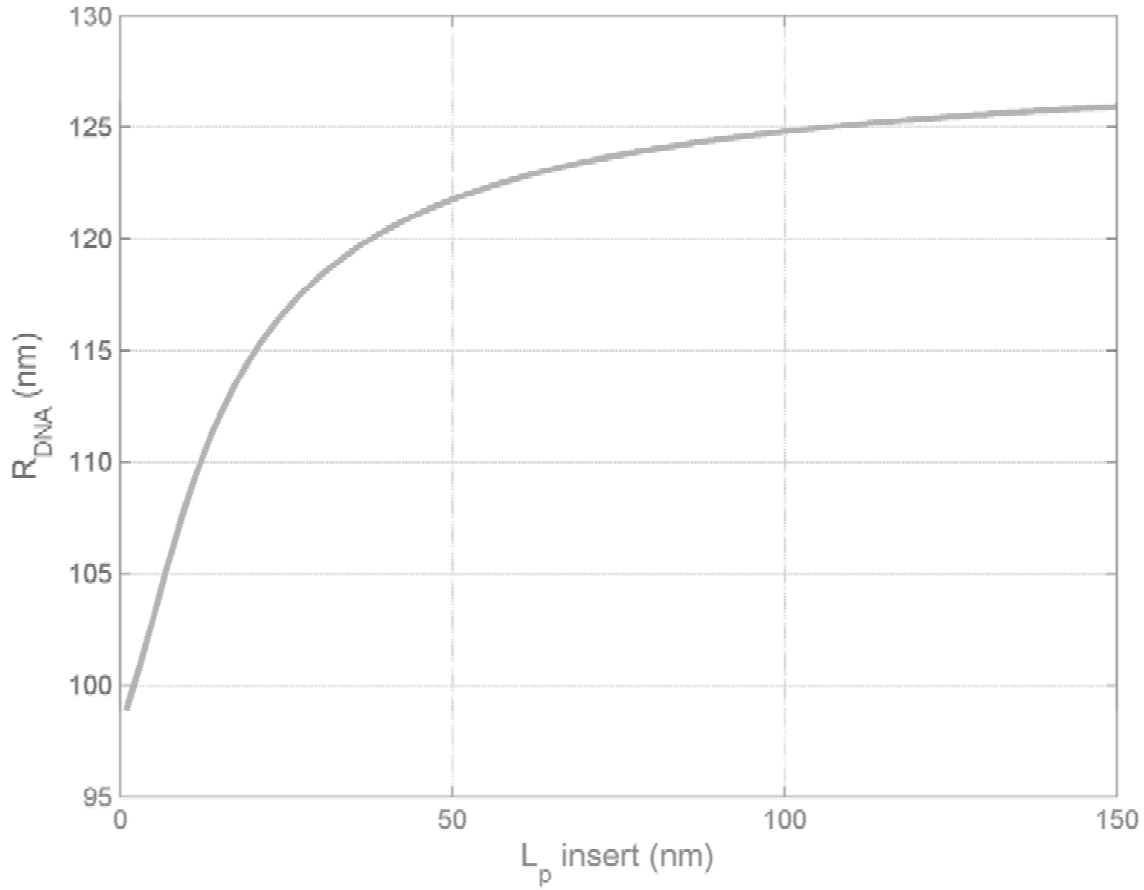
**S3 table:** complement to S2

$$\sqrt{\langle R_{WLC}^2 \rangle} = D[(1 + 0.338 \cos(n\theta_1 - \theta_0)]^{1/2} + \alpha n$$

Fit conditions	R <sub>p</sub>	D (nm)	θ <sub>1</sub> (°)	θ <sub>0</sub> (°)	α (nm)
No global curvature (θ <sub>0</sub> = 0) No local rigidity (α = 0)	150	92 ± 2	15 ± 2		
	155	86 ± 2	17 ± 2		
No local rigidity (α = 0)	150	91 ± 2	19 ± 4	25 ± 19	
	155	95 ± 2	21 ± 4	29 ± 22	
No global curvature (θ <sub>0</sub> = 0) Data fitted from n =0 to 4	150	91 ± 2	17 ± 12		2 ± 4
	155	84 ± 3	22 ± 15		2 ± 4
No global curvature (θ <sub>0</sub> = 0) Data fitted from n =0 to 7	150	92 ± 3	17 ± 12		1 ± 4
	155	86 ± 3	17 ± 15		0 ± 4

**S4 figure: non-linearity of  $R_{DNA}$  with the persistence length of a rigid insert.**

$R_{DNA}$  is computed considering:  $L=575$  bp ,  $l=274$ bp,  $b=88$ bp and with  $L_p=150$  bp outside the 88bp insert. As a result,  $R_{DNA}$  appears to be poorly sensitive to an increase of rigidity of the insert for persistence lengths of the insert  $> 150$  bp.



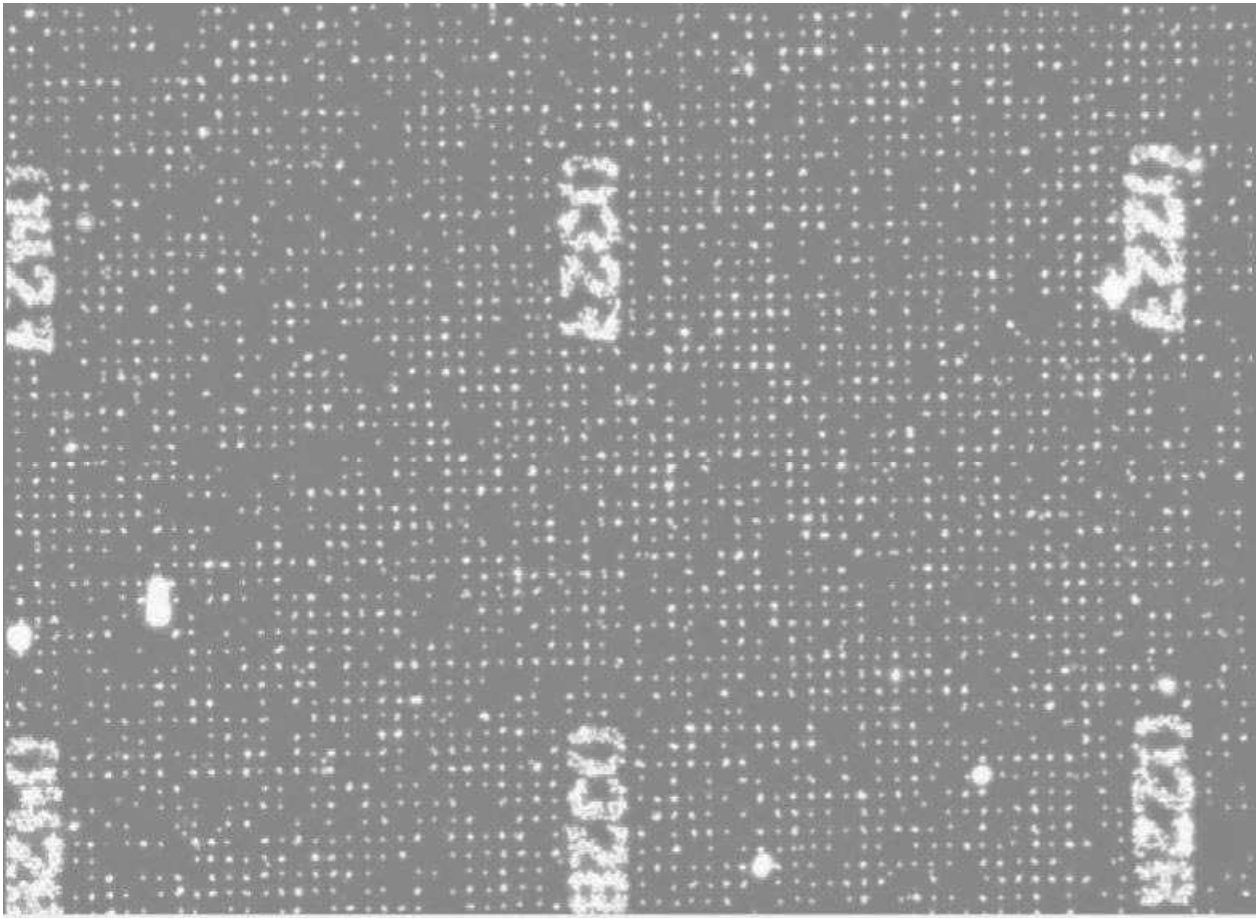


**S5 table**

DNA test	DNA reference	$R_{\text{DNAtest}}$	$R_{\text{DNAref}}$	$\frac{[\langle R_{\text{DNA}}^2 \rangle_{\text{REF}} - \langle R_{\text{DNA}}^2 \rangle_{\text{TEST}}]}{\langle R_{\text{DNA}}^2 \rangle_{\text{REF}}}$	$\theta$ (°)	$\frac{\theta}{n}$ (°)
6An4 <sup>P</sup>	6An0	98.5	105.4	0.127	60	15
6An6 <sup>P</sup>	6An0	95.5	105.4	0.179	73	12
6An7 <sup>P</sup>	6An0	84.1	105.4	0.363	116	17
6An4 <sup>O</sup>	6An0	108.6	105.4	-0.062	-	-

## Video S1

Typical field of Particle-DNA complexes recorded by HT-TPM (area size about 100  $\mu\text{m}$  X150  $\mu\text{m}$ )



# DÉTECTION D'UNE BULLE DE DÉNATURATION EN TPM

## SOMMAIRE

INTRODUCTION . . . . .	255
7.1 L'ADN, UN BIO-POLYMÈRE SENSIBLE À LA TEMPÉRATURE . . . . .	257
7.1.1 Premières observations expérimentales de l'influence de la dénatura- tion sur l'ADN . . . . .	257
7.1.2 Influence de la présence d'une bulle de dénaturation dans le cadre d'un modèle rigide d'ADN . . . . .	258
Idées clés liées à la formation d'une bulle de dénaturation . . . . .	258
Approche selon le modèle rigide . . . . .	258
Influence sur la distance bout-à-bout . . . . .	259
7.1.3 Influence de la température dans le cadre d'un modèle Gaussien d'ADN . . . . .	259
Approche selon le modèle Gaussien . . . . .	259
Influence sur la distance bout-à-bout . . . . .	260
7.1.4 Combinaison des effets liés à la température . . . . .	260
7.2 MODÈLES DE DÉNATURATION . . . . .	261
7.2.1 Modèle simple de dénaturation à l'équilibre . . . . .	261
Modèle de proches voisins (NN) . . . . .	262
Modèle de Poland-Scheraga (PS) . . . . .	262
Validation de la pertinence du modèle PS . . . . .	262
7.2.2 Modèles de dénaturation dynamique . . . . .	264
7.2.3 Approche via notre modèle MC-D . . . . .	264
Estimation de la longueur de persistance d'un double brin . . . . .	264
Estimation de la longueur de persistance d'un simple brin . . . . .	265
Limites du modèle MC-D pour décrire la physique des bulles de dénaturation . . . . .	266
7.2.4 Approche via le modèle Helical (H) . . . . .	266
7.3 STRATÉGIES EXPÉRIMENTALES . . . . .	267
7.3.1 Volet du projet initial relatif à la dynamique de la renaturation . . . . .	267
L'expérience de G. Altan-Bonnet, A. Libchaber et O. Krichевsky . . . . .	267
Estimation de la dynamique des bulles via le modèle <i>Helical</i> . . . . .	268
7.3.2 Volet du projet relatif aux propriétés d'équilibre . . . . .	269
Estimation de l'influence d'une bulle via l'approche Gaussienne . . . . .	269
Estimation de l'influence d'une bulle via notre modèle <b>Stiffer-WLC</b> . . . . .	270
Estimation de l'influence d'une bulle via notre modèle <b>MC-D</b> . . . . .	271

7.3.3	Principe . . . . .	271
	ADNs utilisés . . . . .	272
7.3.4	Estimation de $T_m$ via l'algorithme <i>MeltSim</i> . . . . .	272
7.3.5	Dépendance attendue de $T_m$ en fonction de la force ionique . . . . .	273
7.4	PROCÉDURE EXPÉRIMENTALE . . . . .	273
7.4.1	Contrôle et maintien de la température . . . . .	274
7.4.2	Déroulement d'une expérience . . . . .	275
7.5	INFLUENCE DE LA TEMPÉRATURE SUR L'ÉTAPE DE CORRECTION DU MOYENNAGE VIDÉO . . . . .	276
7.5.1	Contraintes dues à notre limite de résolution temporelle . . . . .	276
7.5.2	Dépendance théorique du coefficient de diffusion en fonction de la viscosité du milieu . . . . .	277
7.5.3	Dépendance de $\tau_{  }$ en fonction de la viscosité du milieu . . . . .	278
7.5.4	Évaluation de la taille de la molécule d'ADN minimale pouvant être utilisée dans cette étude . . . . .	279
	Dépendance de la longueur de persistance en fonction de la température. . . . .	280
7.6	INFLUENCE DE LA TEMPÉRATURE SUR LES PROPRIÉTÉS PHYSIQUES D'UNE MOLÉCULE D'ADN DE 2060 BP . . . . .	282
7.6.1	Dépendance du temps de relaxation en fonction de la température . . . . .	282
7.6.2	Dépendance de la longueur de persistance en fonction de la température. . . . .	283
7.6.3	Estimation de la taille de bulle de dénaturation . . . . .	285
7.7	COMPARAISON AVEC DES DONNÉES TPM DE LA LITTÉRATURE . . . . .	286
7.7.1	Procédure et données TPM extraites de Driessen <i>et al.</i> . . . . .	286
7.7.2	Évaluation de l'influence de la viscosité du milieu dans Driessen <i>et al.</i> . . . . .	288
7.7.3	Estimation du temps de relaxation des données de Driessen <i>et al.</i> . . . . .	289
7.7.4	Estimation des données de Driessen <i>et al.</i> corrigées de l'effet lié à $\tau_{  }$ . . . . .	289
7.7.5	Estimations de $L_p^D$ , extraites des données corrigées du moyennage vidéo . . . . .	290
	CONCLUSION . . . . .	293

## INTRODUCTION

LORS de la division cellulaire, la molécule d'ADN est répliquée à l'identique afin que chacune des deux cellules filles hérite de la même information génétique. Ce processus ainsi que son bon déroulement est donc crucial et vital. La réplication de la double hélice d'ADN, se produisant au cours de la mitose, est initiée par une hélicase qui induit la déshybridation de cette dernière afin de former un *œil de réplication*. Cet œil de réplication correspond donc à l'ouverture locale de la double hélice, via un désappariement des bases complémentaires par la rupture des liaisons hydrogènes entre les bases, séparant l'hélice d'ADN en deux simples brins.

La première image, pas toujours exacte, que nous pourrions associer à ce phénomène de séparation de la double hélice d'ADN, serait que les deux brins qui forment une molécule se séparent à partir de leurs extrémités. Contrairement à cette image, le double brin d'ADN peut spontanément s'ouvrir en son milieu. Ce processus correspond au phénomène de dénaturation. Lorsque cette séparation ou ouverture, se produit sur plusieurs bases successives, ceci occasionne ce que nous nommons une bulle de dénaturation.

*In vivo* ce phénomène peut intervenir à la température physiologique. En effet, les énergies mises en jeu lors des interactions d'empilement dans un même brin ainsi que lors des liaisons hydrogènes entre paires de bases complémentaires, sont de l'ordre de quelques  $k_B T$ , 1.3.2 et 1.2.3. Ces énergies de l'ordre de l'énergie des fluctuations thermiques expliquent que l'ouverture transitoire du double brin soit possible même à température physiologique. Il est donc incontestable que la formation de bulles de dénaturation joue un rôle primordial et principal lors de ce processus biologique.

Ce phénomène de dénaturation qui intervient sur le double brin d'ADN peut être provoqué par des agents chimiques ou physiques, ou causé par un accroissement de la température. Dans cette partie nous nous intéressons à la formation de bulles de dénaturation en fonction de la température.

La définition de la température dite critique, que nous nommons aussi température de fusion ou de dénaturation,  $T_m$ , correspond à la température nécessaire afin que 50 % du double brin d'ADN soit dénaturé, désapparié en deux brins séparés. Les deux brins simples alors formés sont près de cinquante fois moins rigides que l'ADN double brin et ils ont donc un nombre accru de degrés de liberté. Ces deux simples brins peuvent se tordre et se déformer considérablement en prenant de nombreuses configurations qui s'opposent à la recombinaison des brins et à la disparition de la bulle.

Nous est-il possible d'observer et de quantifier l'effet induit par la présence d'une bulle de dénaturation dans l'ADN? Cette variation locale de la structure du double brin influe-t-elle sur le comportement global de l'ADN? Quelles sont les observables qui nous sont accessibles via les expériences TPM? Nous est-il possible d'extraire des informations sur la dynamique de cet événement?

Mais avant cela, quels sont les effets de la température sur l'ADN ? En effet, avant d'atteindre la température de fusion de l'ADN nous allons sonder différentes températures. Or, les propriétés physiques du polymère, telle que la longueur de persistance, sont directement dépendantes de la température. Une variation de la longueur de persistance, paramètre clé pour la description physique des propriétés de l'ADN, va donc directement influencer la mesure de la distance bout-à-bout des complexes ADN/particule. Nous est-il possible d'observer et de quantifier l'effet induit par une variation de température sur les propriétés physiques et mécaniques de l'ADN ?

En résumé cette problématique est double. L'influence de la température conduit à une **modification globale** des propriétés physiques du polymère d'ADN. Une fois la température critique atteinte, l'effet de la température induit un phénomène de dénaturation de la double hélice. Cette dernière étape correspond donc à l'ouverture d'une bulle de dénaturation, **phénomène et défaut local** le long de l'ADN.

L'exploration de l'ensemble de ces phénomènes nous permettrait une compréhension approfondie de leur implication biologique. L'investigation des dépendances des propriétés physiques et mécaniques de la molécule d'ADN en fonction de la température soulève un grand nombre de questions. Nous est-il possible de sonder ces différents régimes de réponse de l'ADN en fonction de la température ? Quelles sont les théories existantes sur la dépendance des propriétés du polymère en fonction de la température ainsi que la dépendance liée à l'ouverture d'une bulle de dénaturation ? Quelles sont les limites de ces théories ? Cela peut-il nous aider à distinguer le modèle le plus pertinent ?

## 7.1 L'ADN, UN BIO-POLYMÈRE SENSIBLE À LA TEMPÉRATURE

Près de la température de dénaturation, une légère augmentation de cette dernière entraîne une perte rapide des multiples interactions faibles maintenant l'ADN sous forme de double hélice. La formation d'une bulle de dénaturation est la conséquence de ce phénomène.

### 7.1.1 Premières observations expérimentales de l'influence de la dénaturation sur l'ADN

L'étude de la déshybridation de la double-hélice d'ADN débute en 1952 grâce aux travaux de Renée Thomas [Tho93]. Ces derniers mettent en exergue le phénomène de dénaturation via les propriétés d'absorption UV de l'ADN. Ces travaux soulignent les différentes contraintes extérieures, thermiques, mécaniques, ioniques, pouvant influencer sur l'ouverture totale ou partielle de la double hélice.

La dénaturation de l'ADN peut, en effet, être étudiée expérimentalement par spectroscopie d'absorption UV. La longueur d'onde utilisée, afin de mesurer la densité optique de l'ADN, est en général de 260 nm. En effet, l'ADN simple brin et la double hélice n'ont pas la même absorption à 260 nm. Le paramètre  $1 - \theta(T)$  correspond à la fraction de paires de bases ouvertes de l'ADN. Ce dernier est défini en fonction de  $Abs(T)$ ,  $Abs_n$  et  $Abs_d$ , l'absorbance UV respectivement associée à la solution d'ADN à la température  $T$ , à l'ADN dans son état double brins et à l'ADN dans son état simple brins, selon l'expression suivante  $1 - \theta(T) = \frac{Abs(T) - Abs_n}{Abs_d - Abs_n}$ . Le coefficient d'absorption optique de la molécule d'ADN augmente d'un facteur 1.4 dès que le double brin se dissocie en deux simples brins. C'est l'effet hyperchrome.

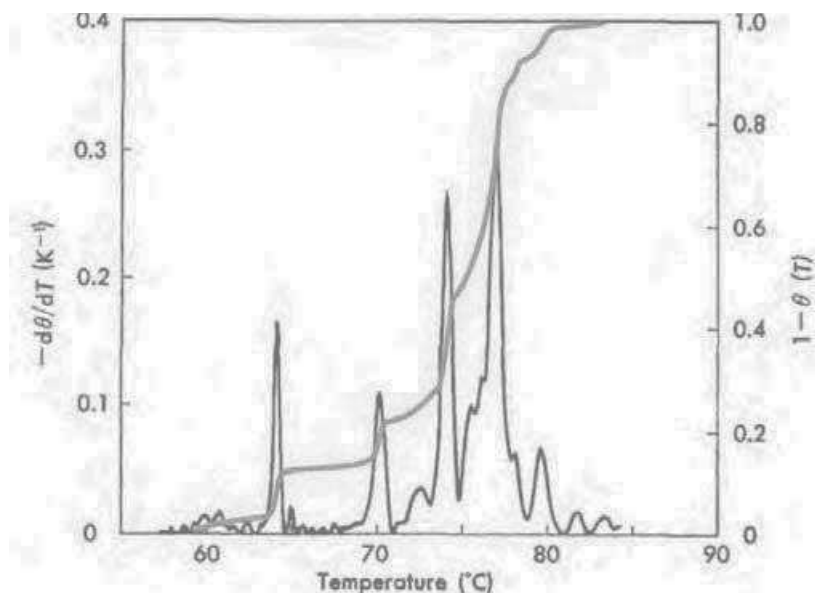


Fig. 7.1 – Courbe de dénaturation thermique d'une séquence d'ADN hétérogène, le plasmide pNT1 de longueur  $L_{ADN} \approx 4$  Kbp. La courbe continue rouge,  $1 - \theta(T)$ , correspond à l'absorbance UV directement proportionnelle au nombre total de paires de bases nouvellement ouvertes, ou profil cumulatif. La courbe continue noire est la dérivée de la courbe précédente. Figure adaptée de [Got83]

La température de dénaturation dépend du désappariement des bases complémentaires à travers la rupture des liaisons hydrogène. L'énergie de liaison du couple de paire de bases **G-C** est évidemment plus importante que celle associée à la liaison **A-T**. Ainsi, la température de fusion de l'ADN est une fonction dépendante de la taille de ce dernier mais surtout de la fraction de paires de base **A-T** et **G-C** de la molécule. Plus la teneur en paire de base **G-C** dans la molécule est importante plus la température de dénaturation augmente.

Les données expérimentales d'absorption UV rassemblées sur la molécule d'ADN montrent sur les courbes un certain nombre de pics abrupts à des différentes températures caractéristiques, Fig. 7.1. Le profil de dénaturation ressemble à une succession de transitions de phases discontinues. Ces pics correspondent à la dénaturation partielle de courtes zones, riches en bases **A** et **T**, de la molécule d'ADN. Nous nous intéressons au phénomène de dénaturation partielle intervenant sur le polymère qui, ici, est souligné par ce type de signature.

Tout ceci conduit les physiciens à s'intéresser à l'étude théorique de la dénaturation thermique. En effet, ce phénomène fut l'objet de nombreuses études réalisées au cours de ces dernières décennies qui ont permis une accumulation de données expérimentales. Ceci permet par conséquent de tester et comparer différents modèles construits afin de décrire les phénomènes associés à la molécule d'ADN.

De plus, l'approche par molécules uniques peut nous permettre d'aller plus loin, dans l'étude fine, en terme de précision, des variations des propriétés physiques de l'ADN, que les mesures d'ensemble réalisées sur ce phénomène de dénaturation. Enfin, l'étude à l'échelle de la molécule unique nous offre une possibilité d'étudier la dynamique individuelle de la molécule d'ADN.

### 7.1.2 Influence de la présence d'une bulle de dénaturation dans le cadre d'un modèle rigide d'ADN

#### Idées clés liées à la formation d'une bulle de dénaturation

En effet, l'élévation de l'énergie thermique provoque l'augmentation des fluctuations et agitations moléculaires. Ceci finit par rompre les forces stabilisant la double hélice d'ADN en provoquant la rupture des liaisons hydrogènes et en déstabilisant les interactions  $\Pi - \Pi$  entre les cycles aromatiques des bases azotées. Les deux simples brins se séparent et s'éloignent l'un de l'autre.

Les simples brins formés sont cinquante fois moins rigides que le double brin. Ainsi, le gain en entropie configurationnelle s'oppose directement à la refermeture de la molécule d'ADN.

#### Approche selon le modèle rigide

Prenons le modèle le plus simple pour décrire un court brin d'ADN, intéressons-nous à l'influence d'une bulle unique de dénaturation présente dans la molécule.

Pour cela considérons l'ADN comme un bâtonnet rigide de longueur  $L < L_p$ . La présence d'une bulle de dénaturation à la position  $i_{bulle}$ , que nous choisirons



centrale afin de faciliter sa prise en compte, est assimilable à la présence d'un coude flexible ou joint libre dans le système, Fig. 7.2. En effet, ceci permet de traduire l'augmentation de l'entropie conformationnelle et du gain en flexibilité associés à la présence d'une zone de simple brin. Ainsi, l'ADN est assimilable à deux sous-bâtonnets rigides, de longueur totale égale à  $L$ , joints entre eux à la position  $i_{bulle} = L/2$ .

### Influence sur la distance bout-à-bout

Ces deux sous-bâtonnets sont considérés comme deux segments statistiquement indépendants.

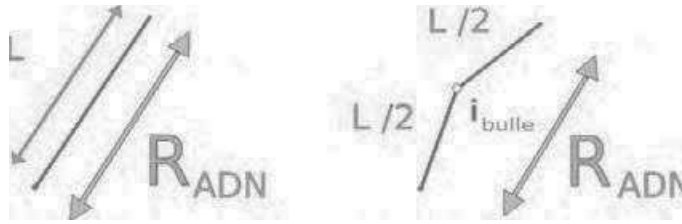


Fig. 7.2 – Représentation de la mesure de la distance bout-à-bout selon le modèle rigide avec et sans présence d'une bulle de dénaturation

La mesure statistique de la distance bout-à-bout n'est alors plus égale à  $R_{ADN} = \sqrt{\langle \mathbf{R}_{ADN}^2 \rangle} = \sqrt{L^2}$  mais vaut maintenant :

$$R_{ADN} = \sqrt{\frac{L^2}{2}} \quad (7.1)$$

La présence de la bulle de dénaturation modifie la mesure de  $\langle \mathbf{R}_{ADN}^2 \rangle$  et donc la mesure de  $\langle \mathbf{R}_{Exp}^2 \rangle$  qu'il nous est possible d'extraire via notre approche TPM de molécule unique.

### 7.1.3 Influence de la température dans le cadre d'un modèle Gaussien d'ADN

Afin de prendre en compte l'effet de la température sur les propriétés physiques de l'ADN il est nécessaire d'intégrer ses caractéristiques élastiques. En effet, la capacité du polymère à se courber dépend de sa longueur de persistance,  $L_p$ , et de son module de courbure,  $\kappa_c$ , ce dernier supposé constant. Comme nous l'avons défini en Eq. 2.3.2, ces grandeurs sont intimement liées via l'équation suivante :

$$\Rightarrow L_p = \frac{\kappa_c}{k_B T} \quad (7.2)$$

#### Approche selon le modèle Gaussien

Utilisons maintenant le modèle de polymère flexible le plus simple, le modèle gaussien. L'effet de la température impactant sur  $L_p$ , se répercute donc sur toutes les grandeurs principales caractérisant le polymère.

L'estimation de la distance bout-à-bout d'un double brin d'ADN, sans présence de bulle de dénaturation, s'exprime alors :

$$R_{ADN} = \sqrt{2LL_p} \quad (7.3)$$

### Influence sur la distance bout-à-bout

En nous basant sur la même description que précédemment intégrons l'effet d'une bulle de dénaturation ainsi que la dépendance de la distance bout-à-bout en fonction de la température pour ce modèle flexible de polymère.

La chaîne est initialement dans l'état totalement hybridée, autrement dit sans bulle de dénaturation. Nous considérons la formation d'une bulle de dénaturation à la température critique  $T_{m,bulle}$ . Cette valeur de température est en dessous de la température de dé-hybridation totale du double brins d'ADN et peut correspondre, comme nous le verrons, à la température de dénaturation d'une séquence locale dans la molécule d'ADN. La mesure de  $R_{ADN}$  est donc modifiée et définie par l'expression suivante :

$$R_{ADN} = \sqrt{\frac{2L}{2} \times \frac{\kappa_c}{k_B T}} \quad (7.4)$$

#### 7.1.4 Combinaison des effets liés à la température

Quelle est l'amplitude, théorique, de la variation de la distance bout-à-bout de l'ADN induite par de la température ?

Pour répondre à cela nous nous intéressons à deux tailles de molécules d'ADN,  $L_{ADN} = 583$  et  $2000$  bp, décrites selon le modèle flexible gaussien et présentant la formation à  $T_{m,bulle} = 70^\circ\text{C}$  d'une bulle de dénaturation en leur milieu, Fig. 7.3.

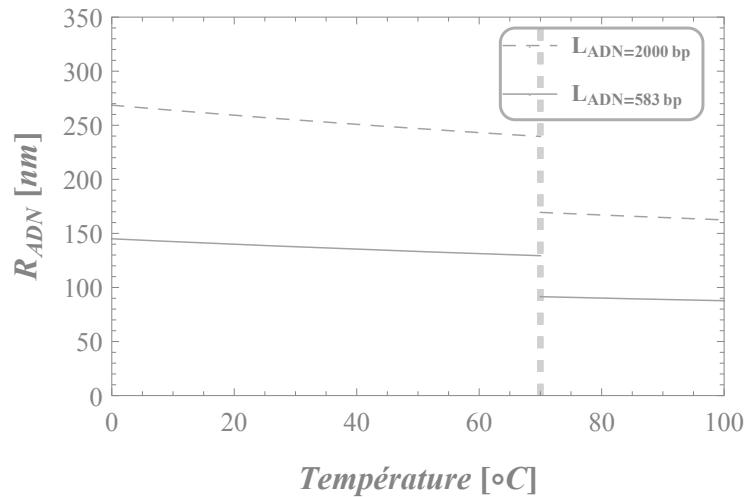


Fig. 7.3 – Variation de la distance bout-à-bout de l'ADN décrit selon le modèle gaussien seul, obtenue pour un ADN de 2000, courbes pointillées, et de 583 bp, courbes pleines, contenant une bulle de dénaturation à partir de  $T_{m,bulle} = 70^\circ\text{C}$  (ligne pointillée verticale) décrite selon un joint libre.

Dans ce cas précis pour  $L_{ADN} = 2000$  bp,  $R_{ADN}$  estimée à  $T_{m,bulle} = 70^\circ\text{C}$  varie de 240 nm à 170 nm suite à l'apparition d'une bulle de dénaturation. Pour  $L_{ADN} = 583$  bp,  $R_{ADN}$  estimée à  $T_m = 70^\circ\text{C}$  varie de 129 nm à 91 nm en présence d'une bulle de dénaturation, Fig. 7.3, soit une variation relative attendue dans le cadre de ce modèle simplifié de  $R_{ADN}$  de  $\sim 30\%$  pour les deux tailles de molécules d'ADN. Nous nous attendons donc à ce que cet effet soit mesurable par HT-TPM.

## 7.2 MODÈLES DE DÉNATURATION

Depuis la première étude expérimentale sur la dénaturation et l'association de brins d'ADN, de nombreux modèles ont été développés afin d'exprimer ce phénomène.

### 7.2.1 Modèle simple de dénaturation à l'équilibre

Deux grandes descriptions standards peuvent être distinguées afin de rendre compte de la dénaturation thermique. Nous allons expliciter les expressions analytiques principales de ces modèles, car nous utiliserons dans la suite un algorithme nommé *MeltSim* basé sur cette approche.

Ces deux modèles mésoscopiques discrets de dénaturation thermique considèrent l'ADN régi par des interactions de type Ising. Autrement dit, tel un modèle d'Ising, le système possède seulement deux états configurationnels possibles :

- un état fermé où les liaisons hydrogène connectant les bases sont présentes
- un état ouvert où ces liaisons sont rompues.

Le principe de ces modèles revient à considérer la dissociation d'un brin d'ADN associé à un couple de base  $N_1 N_2$  en équilibre chimique,  $N_1 N_2 \rightleftharpoons N_1 + N_2$ , avec les états  $N_1$  et  $N_2$  associés aux deux simples brins d'ADN.

Les grandeurs, observables expérimentales, utilisées pour décrire l'équilibre chimique entre ces deux états sont la fraction totale de paires de bases fermées en fonction de la température,  $\theta(T)$ , et le degré d'association,  $\theta_a(T)$ .

L'expression détaillée de la fraction interne de bases fermées,  $\theta_{int}(T)$ , s'exprime de la façon suivante :

$$\theta(T) = \theta_{int, N_1 N_2}(T) \theta_a(T) + \frac{1}{2} (1 - \theta_a(T)) (\theta_{int, N_1}(T) + \theta_{int, N_2}(T)) \quad (7.5)$$

Les deux simples brins sont considérés comme parfaitement complémentaires, sans l'existence d'aucun mésappariement ou d'éventuelles épingles, alors  $\theta_{int, N_1}(T) = \theta_{int, N_2}(T) = 0$ .

La différence de l'énergie libre entre l'état  $N_1 N_2$  et l'état en deux simples brins, est égale à  $\Delta G_0 = \Delta H_0 - T \Delta S_0$ . La loi définissant la constante cinétique de réaction,  $s$ , à l'équilibre chimique s'exprime en fonction de  $\Delta G_0$ , selon l'expression suivante :

$$s = Cte \exp \left( \frac{\Delta G_0}{k_B T} \right) \quad (7.6)$$

La prise en compte des définitions des concentrations respectives associées aux deux états ainsi que la concentration totale, permet d'exprimer le taux d'association en fonction de  $x = \frac{C_0}{C_T} \exp \left( \frac{\Delta G_0}{k_B T} \right)$ , tel que :

$$\theta_a(x) = 1 + x - \sqrt{x(2+x)} \quad (7.7)$$

La température de dénaturation est donc définie par  $\theta_a(T_m) = \frac{1}{2}$ , soit 50% de la molécule est dénaturée. L'utilisation de l'équation précédente permet de définir simplement  $T_m$ , dans le cas où  $\Delta H_0$  et  $\Delta S_0$  sont indépendantes de la température, selon l'expression suivante :

$$T_m = \frac{\Delta H_0}{\Delta S_0 + k_B \log \left( \frac{C_T}{\alpha C_0} \right)} \quad (7.8)$$

Le paramètre  $\alpha$  permet de définir si les brins sont auto-évitants ou non. Les modèles suivants sont basés sur ces lois et définitions.

### Modèle de proches voisins (NN)

Ce modèle, [ZB59, DT62, San98], est utilisé pour décrire quantitativement la dénaturation thermique des oligomères courts, composés de quelques dizaines de paires de bases. C'est un modèle simple et minimal mais efficace dans la description du phénomène de dénaturation sous condition d'une bonne calibration des énergies associées aux différents couples de paires de bases, 1.3.2.

La transition entre les deux états possibles est modélisée via la dépendance de l'énergie libre du couple de bases en fonction de la température. Cette dernière est décrite comme la somme des termes locaux d'énergie d'appariement et d'empilement des paires de bases successives.

### Modèle de Poland-Scheraga (PS)

Ce modèle [PS66], permet de décrire la dénaturation de molécule de polymères plus longue, composée de millier de paires de bases à la différence du modèle NN.

L'énergie libre d'une configuration du système est décrite sur le même principe que le modèle précédent. L'énergie libre est égale, à la fois, à des énergies locales d'association mais aussi à des termes non-locaux, telle que l'énergie de nucléation, associée à la formation d'une bulle de dénaturation, ainsi que l'entropie de conformation associée à la présence de cette bulle. Autrement dit, ce modèle permet de mieux prendre en compte la coopérativité des paires de base. L'état de conformation de chaque paire de bases dépend fortement de l'état de ses voisins.

Ce modèle permet de décrire le phénomène de dénaturation sur des domaines à l'intérieur du double-brin. En revanche, il est nécessaire de souligner que les valeurs numériques des paramètres permettant de décrire l'entropie de la bulle sont des paramètres empiriques. Leur estimation fut améliorée à l'aide de modèles de physique statistique.

L'un des grands avantages de ces deux modèles est le faible temps de calcul associé.

### Validation de la pertinence du modèle PS

Divers modèles, basés sur ce dernier, existent. Certains ont donné lieu à des algorithmes employés afin de reproduire de façon numérique les courbes de dénaturation expérimentale et ce de façon très fidèle.

Le plus connu de ces algorithmes est, sûrement, le programme de bioinformatiques basé sur la résolution du modèle PS nommé *MeltSim* [BBB<sup>+</sup>99], Fig. 7.4 et 7.5.

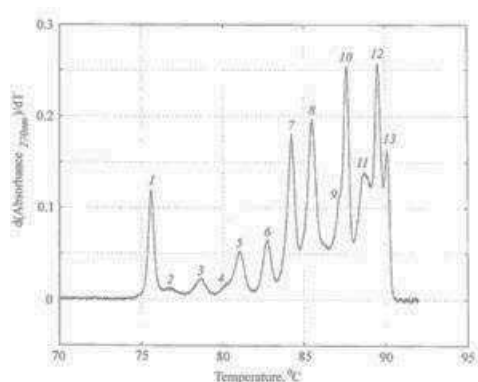


Fig. 7.4 – Courbe expérimentale du profil de dénaturation différentielle de 20  $\mu$ g d'un ADN pN/MCS-13, soit 4664 bp, en présence de 0.0745 mol/L d'ion  $\text{Na}^+$ , le tout suivi à la longueur d'onde de 270 nm. Courbe extraite de [BBB<sup>+</sup>99]

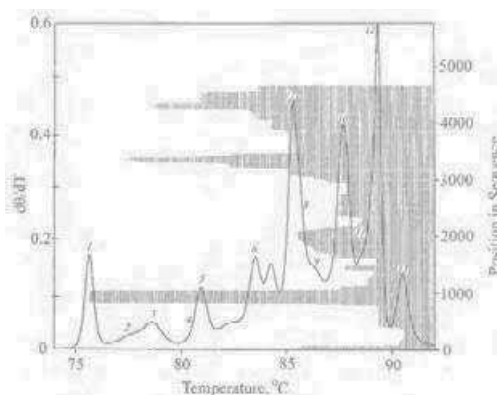


Fig. 7.5 – Courbe calculée du profil de dénaturation différentielle de 20  $\mu$ g d'un ADN pN/MCS-13, soit 4664 bp, Dans les mêmes conditions que l'expérience. Afin de souligner l'accord entre les données expérimentales et numériques, le carte de dénaturation en fonction de la position de la séquence est présentée en grisé sur la droite. Courbe extraite de [BBB<sup>+</sup>99]

L'exemple de la courbe de dénaturation Fig. 7.6 montre bien la bonne concordance entre la mesure de la température *expérimentale* et la valeur estimée via l'algorithme de *MeltSim*.

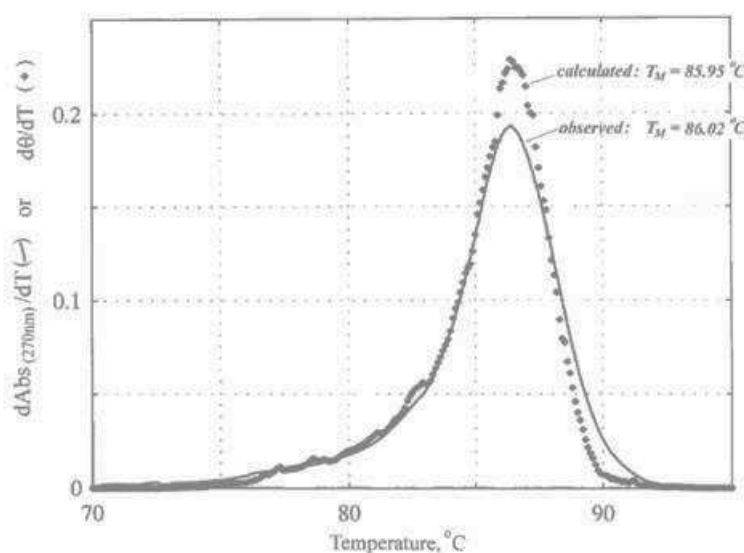


Fig. 7.6 – Courbe expérimentale, ligne continue, et courbe calculée, ligne pointillée, du profil de dénaturation en absorbance UV à la longueur d'onde de 270 nm de l'ADN E-coli, de longueur  $\sim$  4 Kbp, en présence de 0.0745 mol/L d'ions  $\text{Na}^+$ . Courbe extraite de [BBB<sup>+</sup>99]

L'écart entre l'estimation de la température de fusion est inférieur à 0.5°C. Cet algorithme est un véritable outil phénoménologique qui permet de prévoir et quantifier avec précision la température de dénaturation de molécule d'ADN. Comme nous le verrons par la suite, ceci peut nous permettre de tester différentes stratégies expérimentales avant de nous engager dans le vif d'un projet de recherche.

### 7.2.2 Modèles de dénaturation dynamique

Afin de rendre compte de la dynamique de ce phénomène, de nombreux modèles ont été développés. En effet, les modèles précédents, basés sur un jeu de variables à deux états, sont limités dans la description de l'ensemble de phénomènes coopératifs intervenant lors de l'ouverture et/ou fermeture des bulles de dénaturation.

Ces dernières modèles de dénaturation des polymères, modèle de Gao et Prohofsky [GP84], modèle de Dauxois, Peyrad et Bishop [PB89, DPB93] et ou encore le modèle de Barbi, Cocco et Peyrard [BCPR99], sont basés sur une description hamiltonienne continue qui permet une description explicite de la forme des potentiels d'interactions existant entre l'ensemble des bases et paires de bases.

Nous ne rentrerons pas dans les détails de ces différents modèles car ils nous ont paru insuffisants pour appréhender les détails de la dynamique requise ici.

### 7.2.3 Approche via notre modèle MC-D

Notre modèle **MC-D**, précédemment décrit, peut nous permettre d'évaluer l'impact de la présence d'une bulle de dénaturation sur la mesure de la distance bout-à-bout. Rappelons que le détail chimique et atomique des bases azotées n'est pas pris en compte dans notre modèle. La molécule d'ADN est modélisée telle une succession de  $N$  monomères, dont chacun correspond à plusieurs paires de bases successives.

La rigidité de courbure est directement couplée aux degrés de liberté internes représentés par les états des paires de bases, ouvertes ou fermées. Ainsi, une approche permettant la modélisation de la présence d'une bulle de dénaturation consiste à modifier localement la longueur de persistance de la chaîne. Le module de courbure de ce défaut local doit être associé à celui d'un simple brin, ou plus précisément de deux simples brins d'ADN.

Ceci nous permet une approche mésoscopique de la présence d'une bulle de dénaturation sur un ou plusieurs monomères successifs. Seule une variation locale de la longueur de persistance rend compte de la nature déhybridée de la double hélice.

Ce type d'approche nécessite toutefois d'appliquer des valeurs numériques convenables aux longueurs de persistance d'un simple brin,  $L_{p,ss}$  et d'un double brin,  $L_{p,ds}$  d'ADN.

#### Estimation de la longueur de persistance d'un double brin

Comme nous l'avons déjà vu, la valeur de référence, dans les conditions physiologiques de température et de salinité, est communément associée à  $L_{p,ds} = L_{p,Ref} = 50$  nm [Hag88], la valeur communément admise. Nous avons cependant vu dans un chapitre précédant qu'il s'agit seulement d'une première estimation. En effet, notre estimation de  $L_{p,ds}$  dans la gamme de concentration en sel associée aux conditions salines physiologiques pour toutes les tailles d'ADN explorées est de  $L_{p,ds} = 46.6 \pm 4$  nm.

La valeur  $L_{p,Ref}$  est incluse dans notre estimation de  $L_p$  via notre erreur de mesure. Afin d'être le plus *universel* possible, nous resterons avec  $L_{p,ds} = L_{p,Ref}$ .

### Estimation de la longueur de persistance d'un simple brin

La première mesure de la longueur de persistance d'un simple brin d'ADN via une méthode de molécule unique fut réalisée par S. B. Smith, Y. Cui et C. Bustamante en 1996 [SCB96]. La technique de pince optique utilisée permet d'extraire les propriétés physiques de l'ADN, réponse élastique associée à l'application d'une force longitudinale pouvant aller jusqu'à  $\sim 100$  piconewtons. Ainsi étirées les molécules individuelles d'ADN simple brin,  $L_{ADN} \sim 40$  kbp, ont permis d'extraire, dans des conditions salines physiologiques  $C_{Na^+} = 150$  mmol/l, une longueur de persistance de  $L_{p,ss} = 0.75$  nm.

Cette première mesure expérimentale est suivie de près par l'étude menée par B. Tinland, A. Pluen, J. Sturm et G. Weill [TPSW97]. Cette dernière s'intéresse aux mesures des constantes de diffusion de fragments d'ADN allant de 280 bp à 5386 bp par *Fluorescence Recovery After Photobleaching* (FRAP). Cette publication compare et compile ces dernières données avec les données précédentes afin d'étudier la dépendance de  $L_{p,ss}$  en fonction de la force ionique, Fig. 7.7.

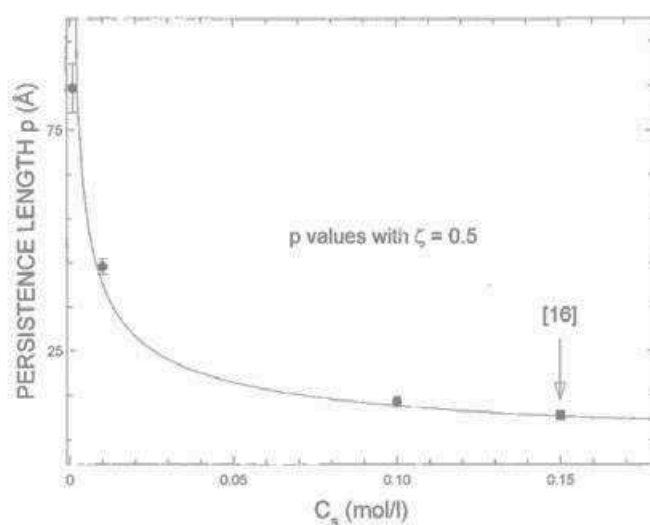


Fig. 7.7 – Courbe expérimentale, de la longueur de persistance d'un ADN simple brin en fonction de la concentration en ion monovalent  $Na^+$  présent en solution. Extrait de [TPSW97], la référence [16] correspond à l'étude [SCB96]

Sur la gamme de force ionique allant des conditions salines aux limites des forts  $I$ , via une formule d'interpolation, l'ensemble des mesures expérimentales est compris entre  $0.8 \text{ nm} < L_{p,ss} < 1.3 \text{ nm}$ , recoupant les données de [SCB96].

En 2001, afin d'étudier la formation en *épingle à cheveux* dans l'ARN et les simples brins d'ADN, A. Montanari et M. Mézard développent un modèle thermodynamique [MM01]. Ce modèle permet d'étendre l'espace des configurations aux structures tertiaires adoptables par l'ADN. Pour cela, il prend en compte l'impact des fluctuations thermiques sur la distance bout-à-bout. La courbe des données numériques décrivant l'allongement de l'ADN simple brin soumis à une force d'étirement est en bon accord avec des mesures expérimentales [MBC00].

Cette comparaison leur permet d'extraire les paramètres décrivant les propriétés physiques de la molécule d'ADN simple brin. Ceci conduit donc à une estimation de la longueur de persistance d'un simple brin égale à  $L_{p,ss} = 1.92$  nm.

Dans notre modèle MC-D, nous imposerons donc  $L_{p,ss} = 2$  nm, au vu des études citées ci-dessus cette valeur correspond à une sur-estimation de  $L_{p,ss}$ . Ceci peut donc nous conduire à une sous-estimation de l'effet sur la distance bout-à-bout induite par la présence d'une bulle de dénaturation. Dans ce cas, les variations extraites de  $R_{||}$  assureront aussi une sous-estimation des valeurs normalement observables sur  $R_{Exp||}$  via notre technique expérimentale.

### Limites du modèle MC-D pour décrire la physique des bulles de dénaturation

Ceci reste cependant une approximation. En effet, la prise en compte de ce défaut local ainsi que l'évaluation de la dynamique associée restent approchées. Les variations des propriétés physiques de l'ADN en présence d'une bulle de dénaturation ne peuvent se résumer simplement à la modification de la longueur de persistance de ce sous-domaine. De plus, la dynamique associée à l'ouverture et la fermeture de cette dernière nécessite une définition précise des potentiels d'interactions mis en jeu.

Afin de répondre avec précision aux questions sur la dynamique ainsi qu'à la mécanique sous-jacente du phénomène d'ouverture et fermeture d'une bulle de dénaturation, un autre modèle mésoscopique a été développé dans l'équipe.

#### 7.2.4 Approche via le modèle Helical (H)

Ce dernier modèle en *gros grains*, Fig. 7.8, se focalise sur la dynamique de séquences courtes d'ADN, allant de la dizaine à 100 bp. La molécule d'ADN est modélisée telle un **homopolymère** constitué de deux chaînes de polymères enroulées l'une sur l'autre afin de rendre compte du caractère hélicoïdal. Chaque simple brin est représenté par une succession de  $N_H$  monomères correspondant à une base azotée, selon le modèle FRC. Le maintien de l'empilement des bases est modélisé via des potentiels harmoniques.

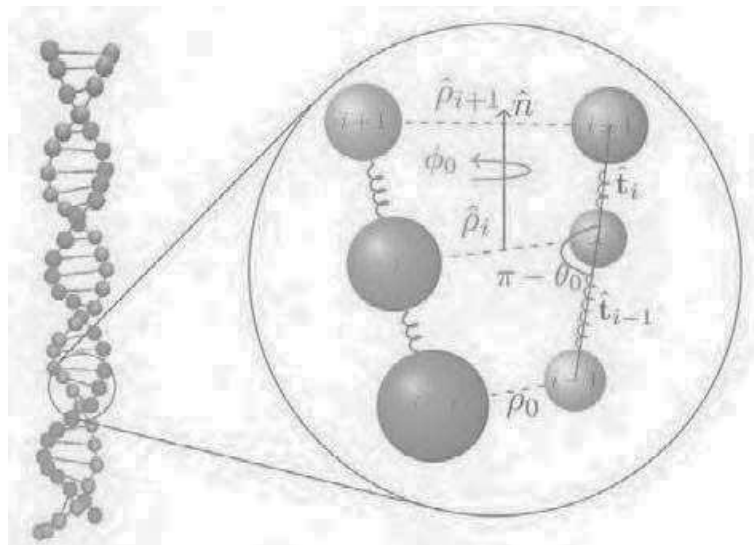


Fig. 7.8 – Représentation de la géométrie du modèle Helical. Extrait de [DDPM13]



Le rayon d'un monomère est donc fixé à  $a_H = 0.17$  nm et la distance entre les bases successives d'un simple brin est égale à  $l_H = 0.357$  nm. L'Hamiltonien associé est constitué de l'énergie de courbure, de l'énergie d'étirement, de l'énergie de torsion, associées à un simple potentiel harmonique, de la chaîne ainsi que de l'énergie associée aux d'interactions hydrogènes, correspondant à un potentiel de Morse.

Soit  $\theta_i$  l'angle déterminé entre deux vecteurs tangents consécutifs le long de chaque simple brin. Ce paramètre est fixé à la valeur  $\theta_0 = 0.41$  rad, chaque simple brin est donc un **FRC**.

Ce modèle permet une approche fine de la dynamique et des mécanismes régissant la fermeture d'une bulle de dénaturation. Le détail des données associées à la dynamique de ce phénomène de dénaturation est décrit dans la suite.

## 7.3 STRATÉGIES EXPÉRIMENTALES

Il reste toutefois délicat d'entreprendre l'étude de l'influence de la présence d'une bulle de dénaturation. En effet, afin de quantifier le plus justement possible les modifications des propriétés physiques engendrées par cette formation, de nombreux points sont à prendre en considération. Tout d'abord, afin d'étudier la formation d'une bulle et une seule, il est nécessaire de connaître, voire d'ajouter la séquence afin de contrôler sa formation. La taille de cette bulle doit être assez importante pour permettre la quantification de son influence sur les observables du système. Toutes ces considérations impactent directement sur les propriétés statiques et dynamiques du système.

Ensuite d'un point de vue purement technique, des difficultés peuvent découler de la nécessité de contrôler la température du milieu. En effet, que ce soit lors de son élévation ou de son maintien, des solutions pratiques doivent être considérées et apportées pour permettre l'exploration d'une large gamme de température.

### 7.3.1 Volet du projet initial relatif à la dynamique de la renaturation

**L'expérience de G. Altan-Bonnet, A. Libchaber et O. Krichevsky**

La première observation de la dynamique d'une seule bulle de dénaturation dans un ADN double brin est rapportée en 2003 par Altan-Bonnet *et al.* [ABLK03].

Les auteurs utilisent une construction spécifique afin de faciliter la formation d'une bulle dans un domaine central constitué de 18 paires de bases **A-T**, le tout flanqué de 7 à 4 paires de bases **G-C** afin d'offrir une plus forte résistance énergétique pour limiter et préserver la séquence d'une déhybridation totale. En effet, cette construction permet de contenir un domaine de *respiration*, phénomène d'ouverture et de fermeture rapide de bulles de dénaturation sur la séquence *Poly(dA-dT)* délimitées et "*scellées*" par des paires de bases **G-C**.

Les mesures sont réalisées via la technique de spectroscopie de corrélation de fluorescence. Pour cela deux stratégies, en terme de disposition du couple fluorophore/inhibiteur de fluorescence (quencheurs), sont utilisées, Fig. 7.9.

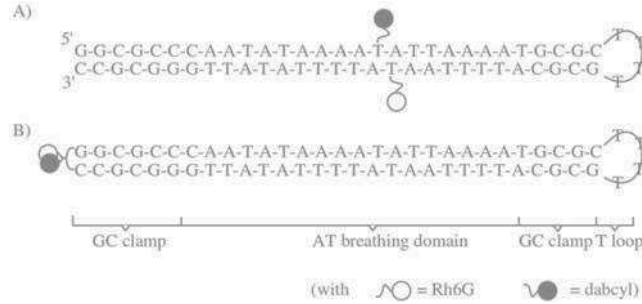


Fig. 7.9 – Représentation des séquences d'ADN utilisées dans l'expérience d'Altan-bonnet [ABLK03]

La première stratégie consiste à placer les marqueurs fluorescents internes à la séquence, dans le domaine *Poly(dA-dT)*. Ainsi, la variation de fluorescence du fluorophore lorsque celui-ci s'éloigne du *quencher*, rend compte de la dénaturation du domaine central. Ceci renseigne donc sur la présence d'une bulle et de sa dynamique de respiration.

La seconde construction dispose les marqueurs fluorescents en bout de séquence à l'extrémité libre de l'ADN. Ainsi, la variation de fluorescence souligne la deshybridation complète de la séquence. L'autre extrémité riche en bases T, reliant les deux simples brins, amène assez de flexibilité pour permettre la formation d'un double brin d'ADN. De plus, cette dernière assure une proximité spatiale des deux simples brins afin que la formation d'un double brin résulte de ces derniers et non de la mise en jeu de simples brins externes à cette séquence. En outre, elle assure la réversibilité de l'expérience de dénaturation-renaturation.

Cette étude a permis de mettre en exergue la formation de bulles et une distribution des tailles fluctuant de 2 à 10 bp présentes dans le domaine de respiration de la construction d'ADN. Ces bulles sont formées au-dessous de la température de fusion  $T_m$  du domaine *Poly(dA-dT)*. L'échelle de temps caractéristique est estimée à  $\sim 50 \mu s$  à  $25^\circ C$ .

### Estimation de la dynamique des bulles via le modèle *Hélical*

Le modèle *Hélical* a permis de montrer que l'énergie de torsion joue un rôle clé dans la dynamique de fermeture de larges bulles de taille allant de 20 bp à 80 bp. Le processus de fermeture se produit en deux grandes étapes, Fig. 7.10.

Premièrement, dans le régime *zipper*, la grande bulle flexible est rapidement refermée par ses deux extrémités jusqu'à atteindre une taille de bulle de  $\sim 10$  bp. La dynamique de cette transition est de l'ordre de  $\tau_{zip} = 130$  ns, associée à la fermeture d'une bulle initiale de taille  $N_{H,bulle} = 100$  bp. La bulle est alors dans un état intermédiaire correspondant à un régime métastable, autrement dit cette structure possède une longue durée de vie, en comparaison au reste du processus.

Deuxièmement intervient la fermeture complète de la bulle. Pour cela plusieurs scénarios sont possibles en fonction de la taille de la séquence  $N_H$ , et principalement du module de torsion,  $\kappa_\Phi$ . Des faibles  $\kappa_\Phi$  induisent un régime *arms diffusion*

*limited* (ADL), la fermeture complète de la bulle est régie par la diffusion des deux bras doubles brins de la séquence et leur alignement. Pour de forts  $\kappa_\Phi$  et de faibles  $N_H$ , le régime *bubble diffusion limited* (BDL), correspond à la diffusion de la bulle le long de la séquence jusqu'à être relaxée par son arrivée à l'une des extrémités de cette dernière. Enfin de forts  $\kappa_\Phi$  et de forts  $N_H$ , conduisent au régime *temperature activated* (TA), correspondant aux mesures expérimentales.

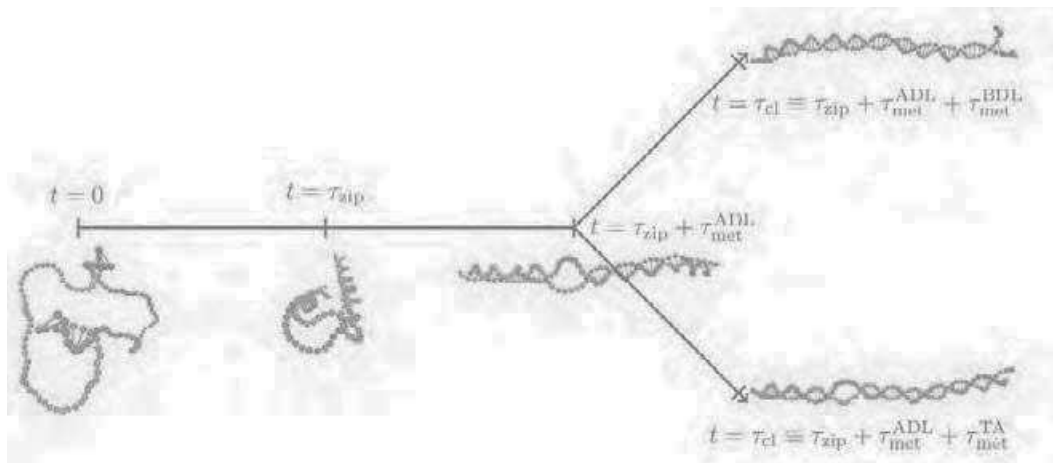


Fig. 7.10 – Représentation des configurations et dynamiques associées du modèle Helical. Extrait de [DDPM13]

Le temps total de fermeture de la bulle est donc égal à la somme des temps moyens de ces différentes étapes. Le temps moyen maximal, correspondant à un régime de fermeture TA, correspondant au régime de fermeture exploré dans l'expérience de Altan-Bonnet *et al.*, d'une bulle de taille  $N_{H,bulle} = 80$  bp est de l'ordre de quelques dizaines de  $\mu s$  et n'est pas significativement dépendant de la taille de la molécule d'ADN.

Actuellement, notre approche expérimentale ne nous permet pas de suivre de telles dynamiques. Pour rappel, notre résolution temporelle actuelle de détection des positions de la particule est de l'ordre de plusieurs dizaines de ms, ceci conduit à une résolution temporelle de  $\sim 2$  s de notre système HT-TPM actuel. Ce volet du projet a donc été momentanément ajourné.

### 7.3.2 Volet du projet relatif aux propriétés d'équilibre

Notre approche simple, décrite précédemment Partie. 7.1.4, estime l'effet de la formation d'une bulle de dénaturation sur la distance bout-à-bout de l'ADN seul. Or, nous cherchons d'abord à déterminer si la chute de la longueur apparente du polymère est détectable par HT-TPM.

#### Estimation de l'influence d'une bulle via l'approche Gaussienne

Reprenons les estimations de la Partie. 7.1.4, afin de calculer la variation attendue sur la distance bout-à-bout des complexes ADN/particule.

Pour cela nous ajoutons analytiquement la contribution de la particule,  $R_{par} = 150$  nm, au mouvement, puis nous estimons la valeur projetée, en 2D, comme obtenue via notre dispositif expérimental, décrit dans la Partie 3.8.2.

La distance bout-à-bout du complexe, comparable à  $R_{Exp||}$  observée par HT-TPM, à  $T_m = 70^\circ\text{C}$  pour  $L_{ADN} = 2000$  bp est de l'ordre de 246 nm pour le cas où il n'y a pas de bulle de dénaturation, et de 194 nm en présence d'une bulle de dénaturation.

Pour  $L_{ADN} = 583$  bp cette distance est de 168 nm sans bulle de dénaturation, et de 147 nm en présence d'une bulle. La variation relative attendue sur  $R_{ADN}$  est de  $\sim 21\%$  pour  $L_{ADN} = 2000$  bp et de l'ordre de  $\sim 13\%$  pour  $L_{ADN} = 583$  bp.

Cette description du comportement de l'ADN en présence d'une bulle de dénaturation, approximée telle un joint libre reliant deux segments gaussiens, ne décrit pas aussi précisément la statistique de conformation des polymères que d'autres modèles, notamment le modèle **WLC**. Afin d'affiner notre évaluation, nous proposons maintenant deux autres approches.

### Estimation de l'influence d'une bulle via notre modèle Stiffer-WLC

Sur les mêmes ADN nous estimons la variation de  $R_{ADN}$  et  $R_{Exp||}$ , en utilisant le modèle **Stiffer-WLC** Partie 6.7.1, à  $T_m = 25^\circ\text{C}$  et en supposant  $L_p = 50.0$  nm .

La résolution de l'équation exacte Eq. 6.7.1, pour l'ADN  $L_{ADN} = 2000$  bp conduit à l'estimation de  $R_{ADN}$  égale à 251 nm. La présence d'une bulle de dénaturation prise en compte en imposant une diminution de la longueur de persistance égale à  $L_p = 2 \times L_{p,ss}$ , sur une séquence locale de 60 paires de bases positionnées en  $l = 773$ , mène à l'estimation  $R_{ADN} = 239$  nm. Après ajout de la contribution de la particule et la prise en compte de la projection en 2D, la mesure obtenue est  $R_{Exp||} = 239$  nm pour l'ADN sans bulle et  $R_{Exp||} = 230$  nm en présence d'une bulle de dénaturation.

Pour l'ADN  $L_{ADN} = 583$  bp cette estimation conduit à  $R_{ADN} = 122$  nm pour l'ADN témoin. La prise en compte de l'ouverture partielle de l'ADN via l'application d'une longueur de persistance de  $L_p = 2 \times L_{p,ss}$  sur une séquence de 60 paires de bases positionnées en  $l = 268$ , mène à l'estimation  $R_{ADN} = 101$  nm.

L'ajout de la contribution de la particule et la prise en compte de la contribution 2D, mène à l'estimation  $R_{Exp||} = 158$  nm pour le polymère sans bulle et  $R_{Exp||} = 148$  nm avec une bulle de dénaturation.

Dans les deux cas, l'effet est donc bien moindre que prévu par notre première estimation.

Ceci conduit, en effet, à une variation absolue observée de  $R_{Exp||}$  de l'ordre de  $\sim 8$  nm pour  $L_{ADN} = 2000$  bp, soit une variation relative de  $\sim 4\%$ . La variation est égale à  $\sim 10$  nm pour  $L_{ADN} = 583$  bp, soit une variation relative de  $\sim 6\%$ .

Ceci souligne bien la difficulté de l'observation et de la quantification de l'effet induit par la présence d'une bulle de dénaturation. En effet, dans le cas attendu où les ADN observés existent dans ces deux états, il sera difficile de distinguer ces deux populations en raison de l'étalement habituel de nos distributions de probabilités qui est de l'ordre de 15 à 20 nm.

### Estimation de l'influence d'une bulle via notre modèle MC-D

Nous réalisons des simulations via le modèle Monte-Carlo dynamique simulant la géométrie de notre système expérimental. Nous simulons la taille d'ADN  $L_{ADN} = 583$  bp, en fixant le nombre de monomères à  $N = (25 + 1)$  auquel est accroché une nanoparticule  $R_{par} = 150$  nm. La taille d'un monomère est donc assimilée à  $a_{bp} = 36$  bp, soit  $\sim 12$  nm.

Tout d'abord, nous réalisons l'estimation, la plus simple correspondant à la mesure de l'influence de la présence d'une bulle de dénaturation sur la mesure de la distance bout-à-bout de l'ADN. C'est à dire que nous allons considérer que le lien entre deux monomères, de 36 bp chacun, est dénaturé. Pour cela nous imposons  $L_p = 2L_{p,ss}$ , intervenant dans l'énergie de courbure mise en jeu entre monomères. Ceci nous permet de conserver le paramètre  $N$  habituel de nos simulations **MC-D**.

La distance bout-à-bout du complexe ADN/particule simulé sans bulle de dénaturation est égale à  $R_{Exp||} = 154$  nm, la modélisation d'une bulle sur 36 bp conduit à une diminution d'amplitude jusqu'à  $R_{Exp||} = 147$  nm. La variation absolue observée de  $R_{Exp||}$  est de  $\sim 9$  nm, soit une variation relative de  $\sim 4.5\%$ .

Ensuite, selon le même principe, nous considérons les liens de deux monomères successifs comme dénaturés. La taille de la bulle de dénaturation est alors de 72 bp, ce qui nous fournit une limite haute de l'amplitude associée à cet effet.

La modélisation d'une bulle sur 72 bp conduit à une diminution de la distance bout-à-bout du complexe ADN/particule simulée de  $R_{Exp||} = 139$  nm. La variation absolue observée de  $R_{Exp||}$  est de  $\sim 15$  nm, soit une variation relative de  $\sim 10\%$ .

Cette estimation, bien que légèrement supérieure à celle obtenue grâce au modèle **Stiffer-WLC**, nous permet toutefois de soupeser de nouveau la difficulté d'observation et de distinction entre les deux états, bulle de dénaturation fermée ou ouverte.

#### 7.3.3 Principe

La stratégie adoptée est donc simple : utiliser pleinement la capacité de dissociation à plus faible température des bases complémentaires **A-T** afin de contraindre et limiter la présence d'une bulle dans ce domaine. Pour cela nous avons fait construire un ADN contenant un cœur *Poly(dA)* de 60 bp. Afin de pouvoir quantifier la différence de comportement de la molécule d'ADN avec bulle et sans bulle, un ADN témoin est aussi construit. Ce dernier possède la même longueur de contour ainsi que la même séquence, sauf sur la région restreinte du cœur *Poly(dA)*. Nous avons étudié la taille de molécules  $L_{ADN} = 583$  bp. Malheureusement, faute de temps nous n'avons pu prospecter expérimentalement l'influence de la température sur une molécule d'ADN de 2000 bp.

Avec notre dispositif expérimental actuel nous ne chercherons pas à extraire la dynamique associée aux phénomènes d'ouverture et de fermeture d'une bulle de dénaturation, ni même à la dynamique associée au régime de respiration d'une bulle dans l'ADN. En effet, cette étude nécessite une vitesse d'acquisition plus rapide que celle de notre caméra actuelle tout en conservant un rapport signal

sur bruit élevé. Le dispositif expérimental est actuellement en cours d'optimisation. Un choix judicieux serait donc de coupler l'utilisation d'une nouvelle caméra avec l'emploi de nanoparticule d'or de 80 nm, possédant une section efficace plus importante.

### ADNs utilisés

L'insert *poly(dA)* est situé au centre de la séquence afin de maximiser son effet. Dans l'ADN témoin, les 60 bp, correspondant à l'absence du *poly(dA)*, sont compensés de part et d'autre des extrémités de la molécule, Fig. 7.11. Ceci permet de limiter les effets que pourrait engendrer cette modification de séquence.

```

TTTTCAGTAAACGGCCAGTTTCATCATAGTAACCTGTATCGTACCAACCTCTTCTTCA
TCTGATCATACCCCAAGAAAGAAATTCCTTTACAGGAGGCTATCAAGTGACCAAGC
GAGAAAACACGGCTTAAACATGGTCCGCTTATCATCAAGCCAGCATTAACGCTCTGAGGA
AACTCAACGAGCTGGGCGGGATGAAACAGCCAGACATCTGTGATGCTTCAAGGALLAGC
TGTATGAGCTTACCGGACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA
AAAAAAGAAACAGCTGGGCGGCTTTGGCTATAGGGGTGAAAGCTTTGACAGAGTACCTCT
GCGAAGAGGCTACAGAGCTCTCTTAAGCGGAGGAGGAGAGCAAGAGCTATTAAGGAGC
ATTAAGGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
CTTATATGAGCTTAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
GTGAAATACCTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG

```

ADN 583 avec un coeur *poly(dA)* de 60 bp

```

TATCTGAGTAAACGGCCAGTTTCATCATAGTAACCTGTATCGTACCAACCTCTTCTTCA
TCTGATCATACCCCAAGAAAGAAATTCCTTTACAGGAGGCTATCAAGTGACCAAGC
GAGAAAACACGGCTTAAACATGGTCCGCTTATCATCAAGCCAGCATTAACGCTCTGAGGA
AACTCAACGAGCTGGGCGGGATGAAACAGCCAGACATCTGTGATGCTTCAAGGALLAGC
TGTATGAGCTTACCGGACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA
AAAAAAGAAACAGCTGGGCGGCTTTGGCTATAGGGGTGAAAGCTTTGACAGAGTACCTCT
GCGAAGAGGCTACAGAGCTCTCTTAAGCGGAGGAGGAGAGCAAGAGCTATTAAGGAGC
ATTAAGGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
CTTATATGAGCTTAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
GTGAAATACCTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG

```

ADN 583 bp

Fig. 7.11 – L'utilité de la représentation des séquences de la molécule d'ADN incluant un *Poly(dA)* et sa molécule témoin réside dans la visualisation rapide, code couleur, de l'insert *poly(dA)* ainsi que leurs différences et similitudes.

### 7.3.4 Estimation de $T_m$ via l'algorithme *MeltSim*

Nous utilisons un programme *MeltSim*, décrit plus haut, afin d'évaluer le profil de dénaturation des molécules d'ADN étudiées en fonction de la séquence de ces dernières. Cet algorithme calcule la valeur de  $T_m$  sur une fenêtre de 50 bp glissante le long de la séquence. Nous estimons le profil de  $T_m$  et du pourcentage de bases G-C des différentes molécules d'ADNs pour une concentration en sel  $C_{[Na^+]} = 155$  mmol/L, afin de se placer dans les conditions physiologiques salines, conditions habituelles de nos expériences HT-TPM.

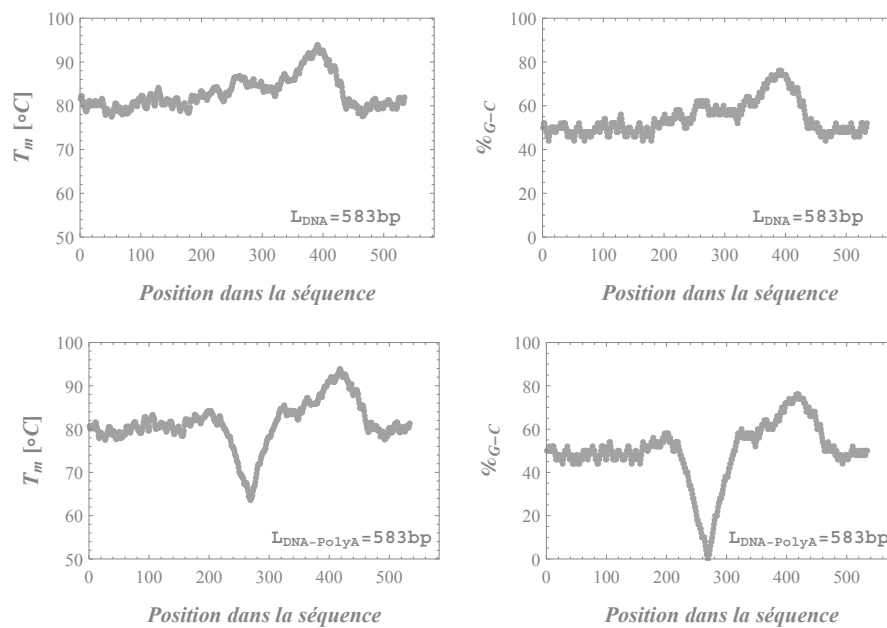


Fig. 7.12 – Variation de la température de dénaturation,  $T_m$ , courbes rouges, et du pourcentage de bases G-C dans la molécule d'ADN,  $\%_{G-C}$ , courbes vertes, calculée en fonction de la séquence de la molécule d'ADN, sur une fenêtre glissante de 50 bp, par l'algorithme *MeltSim*. Les graphiques de la ligne supérieure correspondent aux données de la molécule d'ADN de taille  $L_{ADN} = 583$  bp témoin, la ligne inférieure correspond à  $L_{ADN} = 583$  bp contenant un cœur *Poly(dA)* de 60 bp.

Sur l'ensemble de ces profils, Fig. 7.12, nous pouvons clairement voir l'influence du pourcentage de bases **G-C** de la séquence sur l'estimation de la température de dénaturation. De plus, la présence du *Poly(dA)* de 60 bp induit localement une forte chute de la température de dénaturation à  $T_m \approx 63^\circ\text{C}$  alors que l'ADN témoin présente un minima de température de dénaturation seulement vers  $T_m \sim 75 - 80^\circ\text{C}$ .

### 7.3.5 Dépendance attendue de $T_m$ en fonction de la force ionique

La force ionique influe sur la valeur de la température de dénaturation. Or, pour réduire la température de dénaturation du *poly(dA)* contenue dans la molécule d'ADN de 583 bp il est nécessaire de diminuer fortement la concentration en sel. En effet, pour passer par exemple de  $T_m = 63^\circ\text{C}$  pour  $C_{[\text{Na}^+]} = 155 \text{ mmol/L}$  à  $T_m = 53^\circ\text{C}$  il faut réduire la concentration en ion monovalent jusqu'à  $C_{[\text{Na}^+]} = 5 \text{ mmol/L}$ . Le profil de la diminution de la température de dénaturation est important à très faible force ionique, au-delà sa décroissance est lente comme illustré en Fig. 7.13.

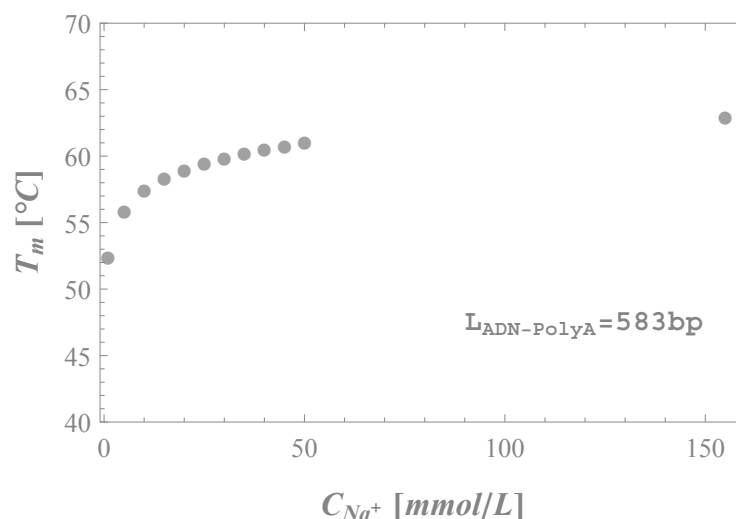


Fig. 7.13 – Variation de la température minimale de dénaturation associée au coeur de *poly(dA)* de 60 bp incluse dans l'ADN  $L_{\text{ADN},\text{Poly}(dA)} = 583 \text{ bp}$  en fonction d'une gamme de concentration en ions  $\text{Na}^+$

Or, comme nous l'avons vu dans le chapitre précédent de telles forces ioniques modifient considérablement les propriétés physiques de la molécule d'ADN, en plus d'augmenter le nombre de décrochages des complexes ADN/particule de la surface d'ancrage. Nous avons donc décidé de travailler à la concentration  $C_{[\text{Na}^+]} = 155 \text{ mmol/L}$ , afin de rester dans les conditions salines physiologiques.

## 7.4 PROCÉDURE EXPÉRIMENTALE

L'ensemble du protocole de préparation de l'échantillon ainsi que les détails du dispositif expérimental correspondent à la description de la Partie 3.1, sauf mention du contraire.

En particulier, la solution tampon utilisée est la solution habituelle, correspondant aux conditions salines physiologiques.

### 7.4.1 Contrôle et maintien de la température

Le contrôle de la température imposée à la chambre microfluidique est assuré par le Peltier *Physitemp* sur une gamme de températures  $T_{\text{Peltier}} \in [-10 : 100]^\circ\text{C}$  en théorie, Partie 3.4.3.

Afin de vérifier la validité de l'application de la température de consigne et nous permettre de quantifier la température réelle de la solution contenue dans la chambre, nous utilisons un dispositif supplémentaire de mesure de la température. Ce dernier est disposé sur le dessus de la chambre micro-fluidique, rappelons que le Peltier chauffe notre échantillon par le dessous. La température de consigne  $T_{\text{consigne}}$ , la température réellement appliquée par le Peltier  $T_{\text{peltier}}$  ainsi que la température évaluée sur l'échantillon  $T_{\text{Cm-M}}$  sont mesurées en début et fin de chaque point de mesure lors de chaque expérience. La variation entre la température réellement appliquée par le Peltier et la température évaluée sur l'échantillon nous permet de calculer la barre d'erreur associée à la température.

Afin de permettre au Peltier de maintenir des températures faibles, essentiellement pour les mesures à  $15^\circ\text{C}$  et  $20^\circ\text{C}$ , un circuit d'eau de refroidissement dont le fonctionnement est régi par la gravité afin d'induire le moins de vibrations possible liées à l'utilisation d'une pompe, est mis en place. Ce circuit est maintenu à débit constant tout le long des expériences pour satisfaire une homogénéité de mesures.

Aux fortes températures la chambre micro-fluidique, bien que juste ouverte sur deux fines ouvertures pour permettre l'injection des complexes ADN/particule, subit un effet d'évaporation. Afin, de pallier à cela des réservoirs de solution tampon, fabriqués en PDMS, sont disposés sur les ouvertures. Pour maximiser l'adhérence de ce dernier l'ensemble de la zone de la chambre est couverte de PMDS comme représenté sur Fig. 7.14. Ce matériau transparent permet de conserver notre bon rapport signal sur bruit.

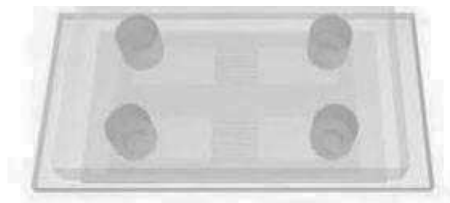


Fig. 7.14 – Représentation de la chambre micro-fluidique que nous venons recouvrir de PDMS afin de fournir des réservoirs de solution tampon limitant l'évaporation de la chambre elle-même

Le déroulement idéal de cette expérience consisterait à effectuer des mesures en montée de températures, puis successivement en descente de températures. Soulignons toutefois que les températures élevées, pouvant atteindre  $80^\circ$  dans notre étude, fragilisent les interactions spécifiques maintenant l'ancrage de l'ADN, streptavidine/biotine, et la fixation de la particule à ce dernier, Dig/Anti-Dig. Nous observons donc un fort taux de décrochage à fort  $T$ , ce qui explique la chute du nombre de trajectoires exploitables des complexes ADN/particule dès que l'échantillon a subi une température de  $60^\circ$  et plus. Le nombre de données acquises est donc limité durant ces expériences en montée/descente de température. Par ailleurs, ces expériences sont très coûteuses en temps, de 3 à 4h pour la montée seulement, pour ces deux raisons, nous les avons donc arrêtées au profit de mesures sur la seule montée de température.



Comme nous venons de le souligner l'augmentation de température induit des décrochages de nos complexes ADN/particule. Ces derniers, libres de diffuser dans toute la chambre peuvent venir perturber notre mesure d'amplitude du mouvement des complexes ADN/particule toujours ancrés sur la surface, comme le montrent les étapes 3.7.9 de l'analyse des données. Des temps de *nettoyage* sont donc nécessaires pour éliminer tout objet non accroché présent en solution. Pour cela tous les deux *pas* de montée en température, nous rinçons extensivement notre échantillon par  $\sim 100 \times V_{canal}$ . Ceci induit des temps supplémentaires afin de permettre à l'ensemble du système de se remettre à température. Suite à différents essais, un temps de 2min à la fin de chaque *nettoyage* garantit, plus que largement, une remise à température de l'ensemble du système ce qui assure, à même valeur de  $T$ , la similarité des mesures avant et après cette étape. Ces étapes de nettoyage peuvent être réalisées sans modification du champ d'observation.

Enfin, la température et l'humidité de la pièce sont relevées en début et fin de chaque expérience afin de contrôler son maintien au cours des mesures. De plus, nous respectons une gamme de valeurs semblables pour l'ensemble des jours de manipulations. Ainsi, la température moyenne doit être de l'ordre de  $20^\circ \pm 1.5^\circ$  lors de chaque expérience. Au cours d'une même expérience, aucune variation supérieure à  $\pm 0.5^\circ$  n'est advenue. De même, l'humidité relative de l'air ambiant est toujours contenue dans la gamme allant de 45% à 70%. Au cours d'une même expérience, la variation moyenne respectée est de l'ordre de  $\pm 1.5\%$ .

### 7.4.2 Déroulement d'une expérience

Une fois l'échantillon prêt, les expériences de dénaturation se déroulent selon les étapes suivantes :

- Une pré-mesure de l'échantillon, sur l'ensemble du dispositif de mesure à température ambiante ou à  $T_{peltier} = 25^\circ\text{C}$ , est effectuée afin de valider la bonne réalisation de notre échantillon *TPM-on-a-chip*.
- Une fois validée, la température de consigne du Peltier est mise à  $T_{consigne} = 15^\circ\text{C}$ . L'ensemble du dispositif de mesure ainsi que la chambre micro-fluidique déjà positionnés sont laissés durant 1h pour assurer la mise à température de l'ensemble. Ce temps est, là aussi, supérieur au temps minimum nécessaire.
- L'ensemble des mesures de température étant stable depuis de nombreuses minutes, nous pouvons commencer le cœur même de l'expérience. Une mesure de 2 minutes est réalisée. Les valeurs des différentes températures sont relevées.
- Une fois l'enregistrement de la vidéo finie, la température de consigne est augmentée d'un *pas* de  $5^\circ\text{C}$ , soit  $T_{consigne,i} = 15^\circ\text{C} + 5^\circ\text{C} \times i$ .
- Une fois la température du Peltier fixée à  $T_{consigne,i} = T_{peltier,i}$  depuis  $\sim 1$  minute, 4 minutes supplémentaires sont laissées à l'ensemble du système pour assurer sa mise à l'équilibre.
- Une mesure d'une durée de 2 minutes est réalisée.
- La température de consigne est augmentée d'un *pas*.
- L'échantillon est *nettoyé*, comme tous les deux *pas* de température.
- 2 minutes plus 4 minutes sont laissées au système pour s'équilibrer.
- Une mesure d'une durée de 2 minutes est réalisée.
- etc..

## 7.5 INFLUENCE DE LA TEMPÉRATURE SUR L'ÉTAPE DE CORRECTION DU MOYENNAGE VIDÉO

Dans un premier temps, nous nous focalisons sur l'analyse des données obtenues pour l'ADN 583 bp sans insert *Poly(dA)* de 60 bp.

### 7.5.1 Contraintes dues à notre limite de résolution temporelle

Tout d'abord, intéressons-nous à la fluctuation de la variable brute que nous mesurons directement via notre technique HT-TPM, c'est à dire non corrigée par le temps de relaxation du système.

Sur la gamme de température prospectée, nous observons une décroissance significative de la distance bout-à-bout des complexes ADN/particule en fonction de la température, Fig. 7.15.

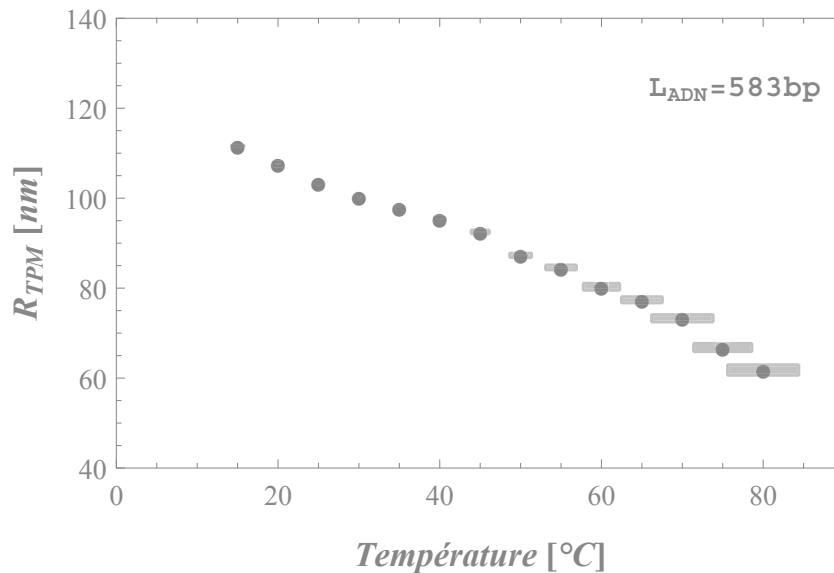


Fig. 7.15 – Variation de la distance bout-à-bout des complexes ADN/particule en fonction de la température pour  $L_{ADN} = 583$  bp témoin (sans l'insert *Poly(dA)* de 60 bp). Les barres d'erreur en  $R_{TPM}$  sont calculées par estimation bootstraps et  $2\sigma_B$  sont représentés. L'erreur en  $T$  correspond à l'écart entre la mesure de  $T_{peltier}$  et  $T_{sonde}$ , là aussi,  $2\sigma_T$  sont représentés

L'extraction du temps de relaxation moyen en fonction de chaque point de température, met vite en lumière certaines difficultés. En effet,  $\tau_{||}$  passe très vite en dessous de la limite de résolution temporelle associée à notre dispositif expérimental qui est égale à 13 ms, Fig. 7.16.

Il est donc impossible d'exploiter ces données pour toutes les températures supérieures à  $T = 35^\circ\text{C}$ . De plus, à  $T = 35^\circ\text{C}$ , un grand nombre de trajectoires, supérieur à 10%, possède un temps de relaxation moyen inférieur à notre limite de résolution temporelle égale à  $\tau_{||} = 13$  ms.

Sur le peu de données exploitables des mesures de la distance bout-à-bout de l'ADN seul corrigées du moyennage vidéo, nous n'observons pas de dépendance significative en fonction de la température, Fig. 7.16

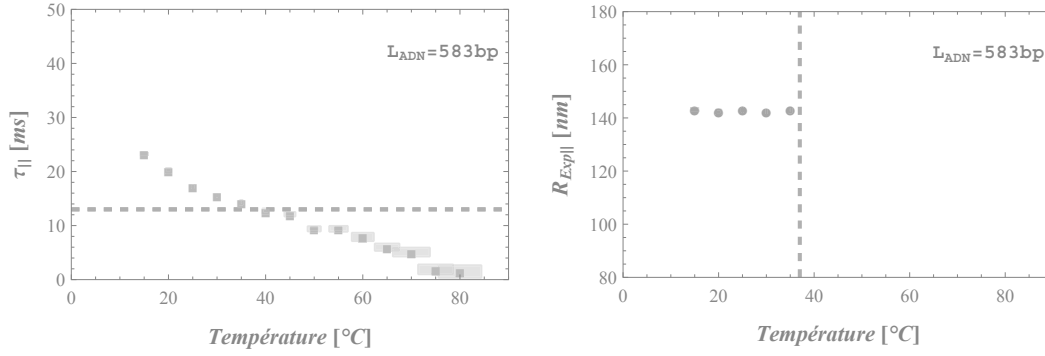


Fig. 7.16 – Variation du temps de relaxation des complexes ADN/particule  $\tau_{||}$ , à gauche, variation de la distance bout-à-bout des complexes ADN/particule  $R_{Exp||}$ , à droite, le tout en fonction de la température pour  $L_{ADN} = 583 \text{ bp}$  témoin

### 7.5.2 Dépendance théorique du coefficient de diffusion en fonction de la viscosité du milieu

Le temps de relaxation des complexes ADN/particule rend compte de la diffusion des systèmes dans le fluide, soumise au mouvement brownien. Or, le coefficient de diffusion d'un objet dans un fluide est proportionnel à la température du milieu et inversement proportionnel à la viscosité de ce dernier,  $\eta$ , selon l'équation suivante :

$$D_{iff} = \frac{k_B T}{6\pi\eta R} \quad (7.9)$$

où  $R$  est, par exemple, le rayon d'une sphère. Dans notre cas, la dynamique de diffusion est dominée par le coefficient de diffusion de la nanoparticule, plus lent que celui de l'ADN comme nous l'avons vu en 3.1.

Basée sur des mesures expérimentales, une loi empirique permet d'exprimer la viscosité de l'eau en fonction de la température selon l'équation suivante [Hay14] :

$$\eta = 2.414 \times 10^{-5} \times 10^{\frac{247.6}{T[K]-140}} \quad (7.10)$$

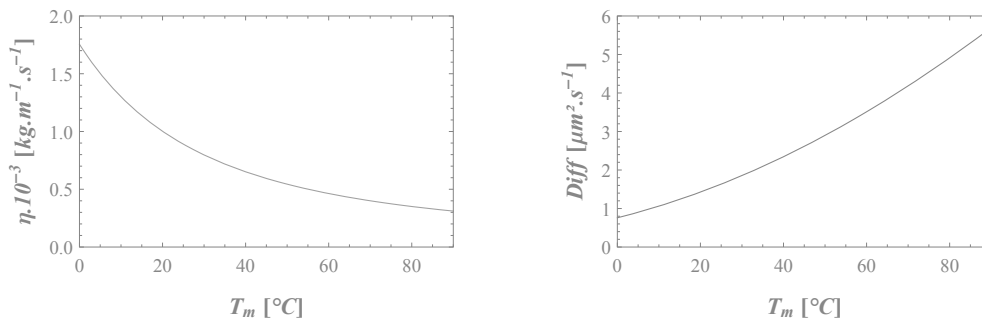


Fig. 7.17 – Variation de la viscosité,  $\eta$ , d'un milieu aqueux en fonction de la température, à gauche. Variation du coefficient de diffusion de la nanoparticule  $R_{par} = 150 \text{ nm}$ , utilisée durant nos expériences, à droite

7.5.3 Dépendance de  $\tau_{||}$  en fonction de la viscosité du milieu

Sur le même principe nous pouvons évaluer quantitativement la dépendance de  $F_{TPM} = D_{diff} \times \tau_{||}$  associée à nos données, dans l’approximation où  $\eta_{TPM} = \eta$  paramètre intervenant dans le calcul de  $D_{diff}$  pour  $R_{par} = 150\text{ nm}$ , Fig. 7.18.

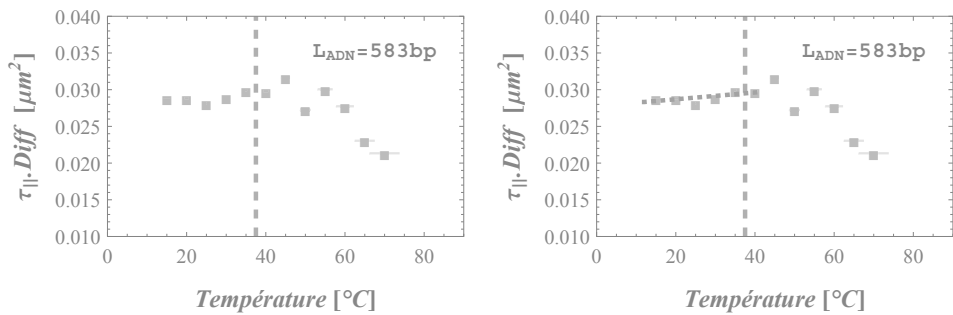


Fig. 7.18 – Pour  $L_{ADN} = 583\text{ bp}$  témoin, variation de  $F_{TPM}$ , multiplication du temps de relaxation de nos complexes ADN/particule,  $\tau_{||}$  par le coefficient de diffusion  $D_{diff}$  d’une particule de rayon  $R_{par} = 150\text{ nm}$ , en fonction de la température. Sur le graphique de gauche la droite violette est une régression linéaire des données respectant  $\tau_{||} > 13\text{ ms}$ .

La variation de  $F_{TPM}$  en fonction de la température est, comme attendu, très faible. En regardant de plus près, il apparaît une faible augmentation, évaluée selon une fonction affine, sur les données associées à  $\tau_{||} > 13\text{ ms}$ . En effet les derniers points, où la dépendance de  $F_{TPM}$  est inversée montrent un comportement singulier de  $\tau_{||}$ , sûrement dû à notre erreur sur l’estimation du temps de relaxation.

	Expression du fit : $F_{TPM} = A.T + B$	
	$A, \mu m^2.K^{-1}$	$B, \mu m^2$
$L_{ADN} = 583\text{ bp}$	$46.2 \times 10^{-6}$	$27.8 \times 10^{-3}$

L’ensemble de ces données,  $R_{TPM}$ ,  $\tau_{||}$  et  $R_{Exp||}$ , ne nous permet pas de conclure quant à l’influence de la température, ni de la présence de bulle de dénaturation, sur les propriétés physique d’un ADN de 583 bp. En revanche cela nous a permis d’observer l’influence de la température sur la viscosité du milieu qui se répercute sur le temps de relaxation des complexes ADN/particule. Cette décroissance de la viscosité du milieu conduit à ce que nous atteignons notre limite de résolution temporelle à températures faibles, soit 35°C pour l’ADN de 583 bp.

Forts de cette première observation, nous avons décidé de tester les molécules d’ADN de différentes longueurs précédemment utilisées. Ceci va nous permettre de sonder l’influence de la température sur  $\tau_{||}$  pour différentes tailles d’ADN. Ainsi, nous pourrons définir les limites de  $L_{ADN}$  pour lesquelles il est possible de prospecter l’influence de la température sur une gamme étendue, allant de 15°C à 80°C, sans dépasser notre limite de résolution temporelle.

Notons toutefois que des essais furent aussi réalisés afin de ralentir la dynamique des systèmes par l'ajout d'un fort pourcentage de glycerol en solution. Or, pour obtenir un temps de relaxation convenable à  $T = 80^\circ\text{C}$ , 50% de glycérol est nécessaire, soit une augmentation de  $\tau_{||}$  d'un facteur 6. Ces essais ne furent pas convaincants, dû au peu de données acquises et au coût de temps d'expérience très important, et donc laissés de côté.

#### 7.5.4 Évaluation de la taille de la molécule d'ADN minimale pouvant être utilisée dans cette étude

Intéressons-nous à la molécule d'ADN  $L_{ADN} = 1201$  bp. Cette taille de molécule nous permet de définir les tailles limites de  $L_{ADN}$  qui nous permettent de débiter notre investigation de l'influence de la température

Les données TPM brutes montrent toujours une décroissance en fonction de la température, Fig. 7.19.

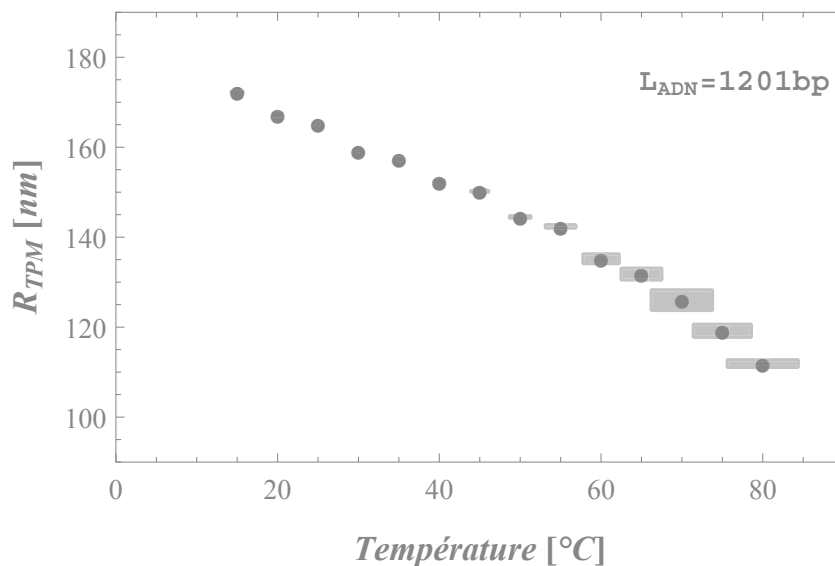


Fig. 7.19 – Variation de la distance bout-à-bout brute des complexes ADN/particule en fonction de la température pour  $L_{ADN} = 1201$  bp

Pour cette longueur d'ADN, la limite temporelle associée à notre dispositif expérimental est atteinte aux alentours de  $T = 60^\circ\text{C}$ , Fig. 7.20.

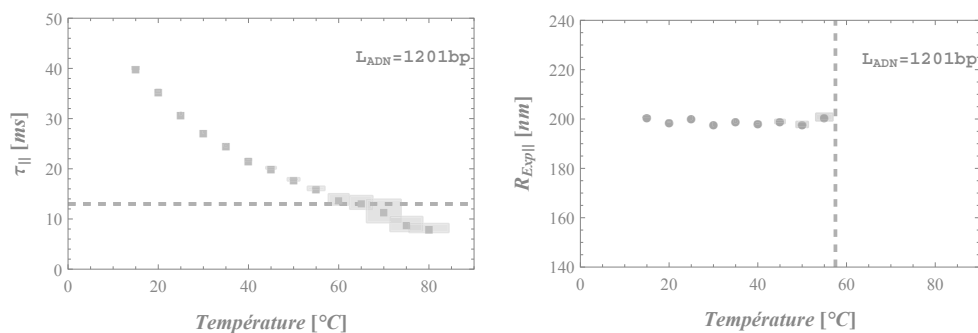


Fig. 7.20 – Variation du temps de relaxation des complexes ADN/particule, à droite, variation de  $R_{Exp}$  [nm], à gauche, le tout en fonction de la température.

Nous n'observons pas de dépendance significative en fonction de la température, sur les données exploitables pour  $L_{ADN} = 1201$  bp. Ces dernières s'étendent sur la gamme de températures allant de  $T = 15^\circ\text{C}$  à  $T = 60^\circ\text{C}$ , Fig. 7.21.

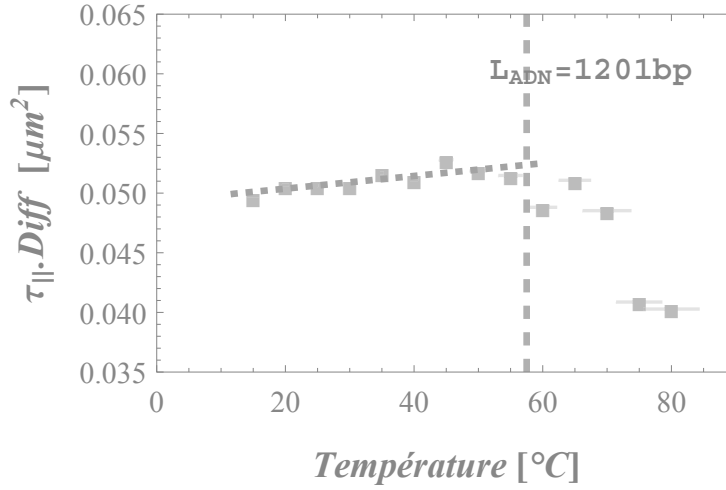


Fig. 7.21 – Variation de  $F_{TPM}$ , multiplication de  $\tau_{||}$  par  $D_{diff}$ , en fonction de la température, pour  $L_{ADN} = 1201$  bp, ainsi que son ajustement selon une fonction affine, droite violette, sur les données respectant  $\tau_{||} > 13$  ms.

L'évaluation quantitative de la dépendance de  $F_{TPM} = D_{diff} \times \tau_{||}$  associée aux données de l'ADN de 1201 bp est comparable à l'ajustement des données associées à l'ADN sans *Poly(dA)* décrit précédemment.

	Expression du fit : $F_{TPM} = A.T + B$	
	$A, \mu\text{m}^2.K^{-1}$	$B, \mu\text{m}^2$
$L_{ADN} = 583$ bp	$46.2 \times 10^{-6}$	$27.8 \times 10^{-3}$
$L_{ADN} = 1201$ bp	$53.2 \times 10^{-6}$	$49.3 \times 10^{-3}$

Le calcul de la fonction  $F_{TPM}$  semble être un outil intéressant qui peut nous permettre de mettre en lumière les limites associées à notre résolution temporelle ainsi que les erreurs induites par la correction de  $R_{TPM}$  par l'effet de moyennage vidéo. En effet, lorsque la limite de résolution temporelle est atteinte, la dépendance en la température de  $F_{TPM}$  semble brusquement inversée. Ceci signifie que  $F_{TPM}$  nous permet de sonder et visualiser un phénomène de comportement singulier du temps de corrélation.

### Dépendance de la longueur de persistance en fonction de la température.

Les lois de dépendance de  $R_{Exp||}$  en fonction de  $L_p$ , extraites grâce aux simulations via le modèle d'échantillonnage exact et présentées dans la partie *Influence de la force ionique* 5, nous permettent d'extraire  $L_p$ , selon la méthode présentée en 5.2.3.

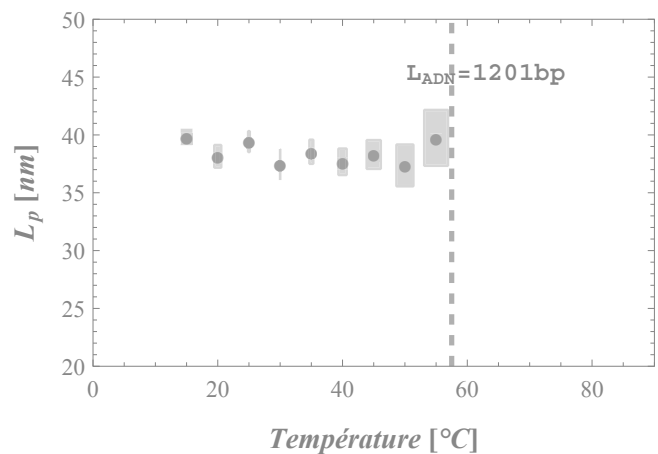


Fig. 7.22 – Variation de  $L_p$  en fonction de la température, pour  $L_{ADN} = 1201$  bp

Nous évaluons la dépendance de  $L_p$  en fonction de la température par un ajustement via une fonction affine  $L_p = \frac{A}{T} + B$  afin de tester la validité de la loi suivante, déjà présentée en Eq. 7.11 :

$$\Rightarrow L_p = \frac{\kappa_c}{k_B T} \quad (7.11)$$

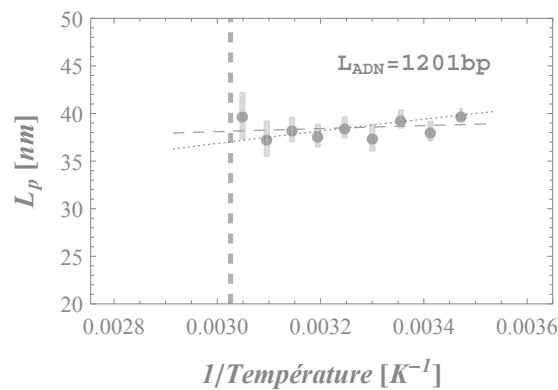


Fig. 7.23 – Variation de  $L_p$  en fonction de  $1/T$ , pour  $L_{ADN} = 1201$  bp. Droite noire ajustement avec 1 paramètre libre, droite rouge avec 2

La dépendance  $L_p$  en fonction de  $T$  est linéaire et présente une décroissance lente, Fig. 7.22 et 7.23.

	Expression du fit : $L_p = \frac{\kappa_c}{k_B T} + L_p^c$	
	$\kappa_c$ , en J.nm	$L_p^c$ , en nm
$L_{ADN} = 1201$ bp	$3.1 \times 10^{-19}$	-
	$0.8 \times 10^{-19}$	29.0

Afin de prospector l'influence de la températures sur les propriétés physiques et mécaniques de la molécules d'ADN pour la gamme de température allant de  $T = 15^\circ\text{C}$  et  $T = 80^\circ\text{C}$ , il est nécessaire d'utiliser une taille de molécule d'ADN supérieure à 1201 bp. En effet, cette dernière, au vu de notre limite de résolution temporelle et de l'effet de la viscosité sur le temps de relaxation des complexes ADN/particule, nous permet d'explorer seulement la gamme de  $T$  allant de  $T = 15^\circ\text{C}$  et  $T = 60^\circ\text{C}$ .

## 7.6 INFLUENCE DE LA TEMPÉRATURE SUR LES PROPRIÉTÉS PHYSIQUES D'UNE MOLÉCULE D'ADN DE 2060 BP

Les données TPM brutes indiquent toujours une dépendance en fonction de la température, Fig. 7.24.

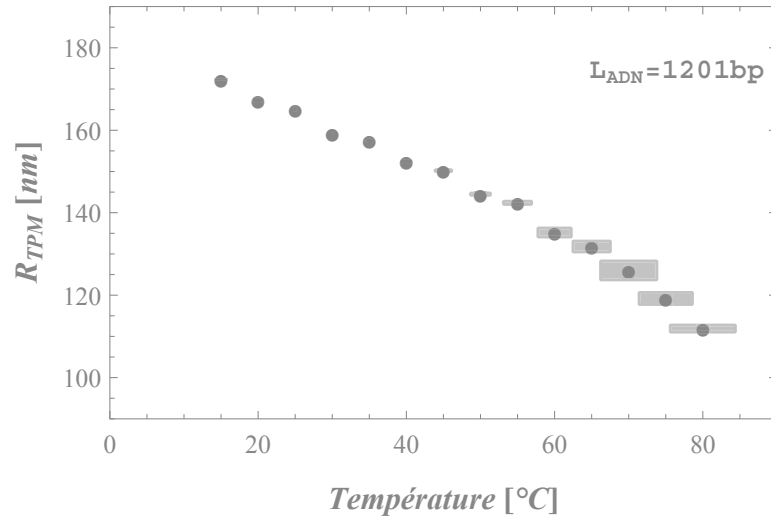


Fig. 7.24 – Variation de la distance bout-à-bout des complexes ADN/particule en fonction de la température pour  $L_{ADN} = 2060$  bp

Pour cette longueur d'ADN, la limite temporelle associée à notre dispositif expérimental n'est pas atteinte sur la gamme de températures prospectées, Fig. 7.25. Il est donc possible de corriger  $R_{TPM}$  de l'effet lié au moyennage vidéo.

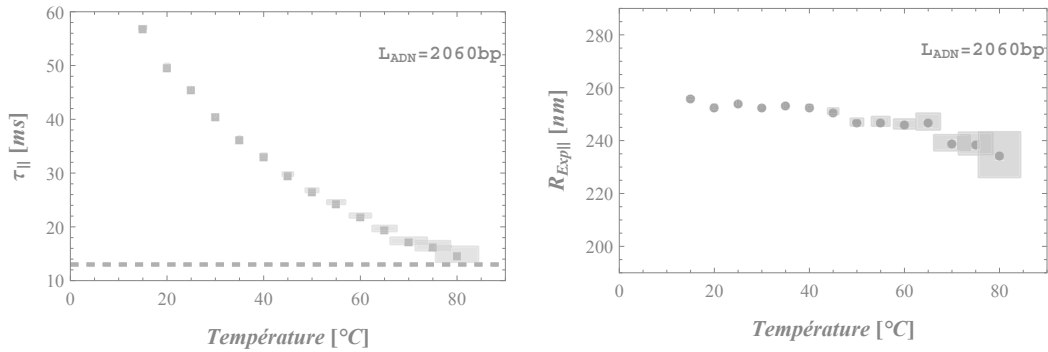


Fig. 7.25 – Variation du temps de relaxation des complexes ADN/particule, à gauche, variation de  $R_{Exp||}$ , à droite, le tout en fonction de la température.

La variation de  $R_{Exp||}$  en fonction de la température, admet une lente décroissance. De plus, entre les températures  $T = 65^\circ\text{C}$  et  $T = 70^\circ\text{C}$ , un décrochage de  $R_{Exp||}$  est observé, Fig. 7.25.

### 7.6.1 Dépendance du temps de relaxation en fonction de la température

L'évaluation quantitative de la dépendance de  $F_{TPM} = D_{iff} \times \tau_{||}$  suit le même profil de croissance qu'obtenu pour les ADN précédents, mais seulement sur la gamme de températures inférieures à  $T < 60^\circ\text{C}$ , Fig. 7.26.



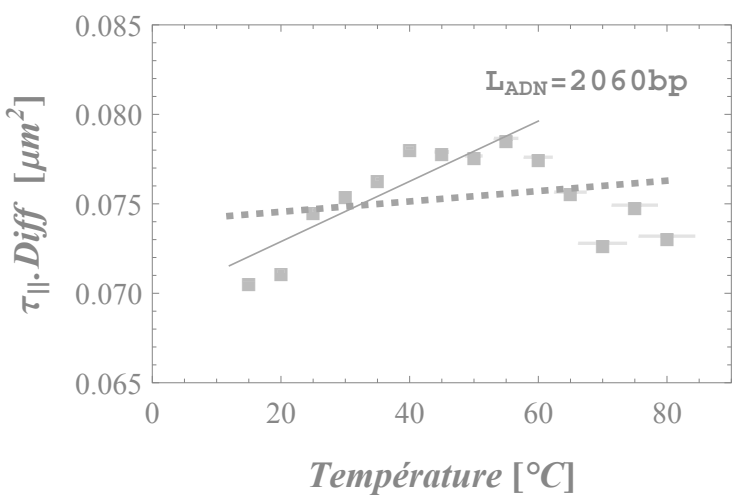


Fig. 7.26 – Variation de  $F_{TPM}$ , multiplication de  $\tau_{||}$  par  $D_{iff_{par}}$ , en fonction de la température, pour  $L_{ADN} = 2060$  bp, ainsi que son ajustement selon une fonction affine, droite violette épaisse, sur toutes les données, droite violette fine, sur les données où  $T < 60^{\circ}\text{C}$ .

Un nouvel ajustement via une fonction affine est donc réalisé sur cette gamme de températures restreintes, seuls les valeurs de ce dernier sont renseignées dans le tableau.

	Expression du fit : $F_{TPM} = A.T + B$	
	$A, \mu m^2.K^{-1}$	$B, \mu m^2$
$L_{ADN} = 583$ bp	$46.2 \times 10^{-6}$	$27.8 \times 10^{-3}$
$L_{ADN} = 1201$ bp	$53.2 \times 10^{-6}$	$49.3 \times 10^{-3}$
$L_{ADN} = 2060$ bp	$168.3 \times 10^{-6}$	$69.5 \times 10^{-3}$

L'inversion de pente observée de  $F_{TPM}$  intervient lorsque le temps de relaxation des complexes ADN/particule est mesuré dans la gamme  $\tau_{||} \in [27, 15]$  ms. Ceci n'est donc pas associé aux limites de notre résolution temporelle et de l'erreur induite par la correction de  $R_{TPM}$ . Ici, ceci souligne donc, plutôt, un phénomène de dépendance singulière du temps de corrélation.

De plus, ce phénomène apparaît dans la gamme de températures correspondante à la chute observée sur les valeurs d'amplitude de  $R_{Exp||}$ . Une dénaturation partielle pourrait expliquer ce phénomène. En effet, la formation d'une bulle de dénaturation rend l'ADN plus flexible, ce qui doit influencer sur la mesure de son temps de relaxation et induire une diminution de ce dernier.

### 7.6.2 Dépendance de la longueur de persistance en fonction de la température.

À partir de la loi de dépendance de  $R_{Exp||}$  en fonction de  $L_p$  obtenue grâce aux simulations EE, 5.2.3, nous extrayons de nouveau  $L_p$ , Fig. 7.27.

Nous évaluons, ici aussi, la dépendance de  $L_p$  en fonction de la température via une fonction affine  $L_p = \frac{\kappa_c}{k_B T} + L_p^c$  afin de tester la validité de la loi de l'Eq. 7.11.

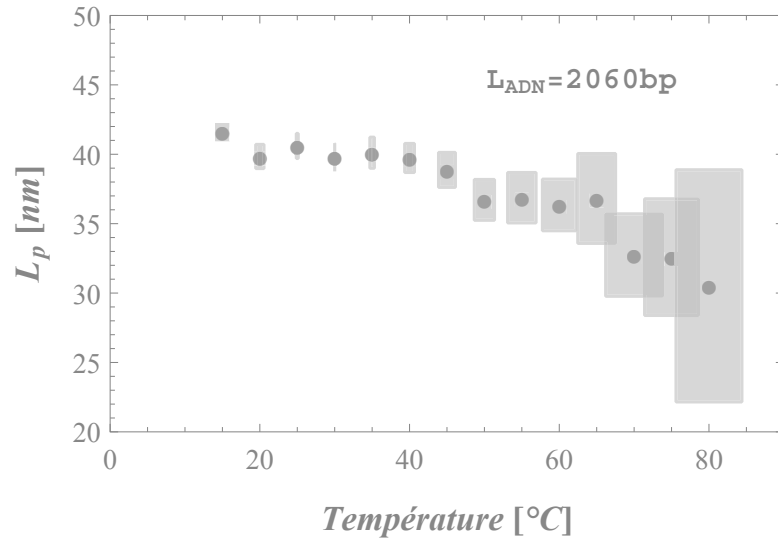


Fig. 7.27 – Variation de  $L_p$  en fonction de la température, pour  $L_{ADN} = 2060$  bp

Pour cela nous réalisons un ajustement sur la gamme entière de températures, la qualité médiocre de celui-ci nous amène à considérer deux régimes de dépendance de  $L_p$  en fonction  $T$ . Pour cela nous réalisons un ajustement de cette dépendance sur les gammes restreintes  $T \in [15 : 65]^\circ\text{C}$  et  $T \in [70 : 80]^\circ\text{C}$ , Fig. 7.28.

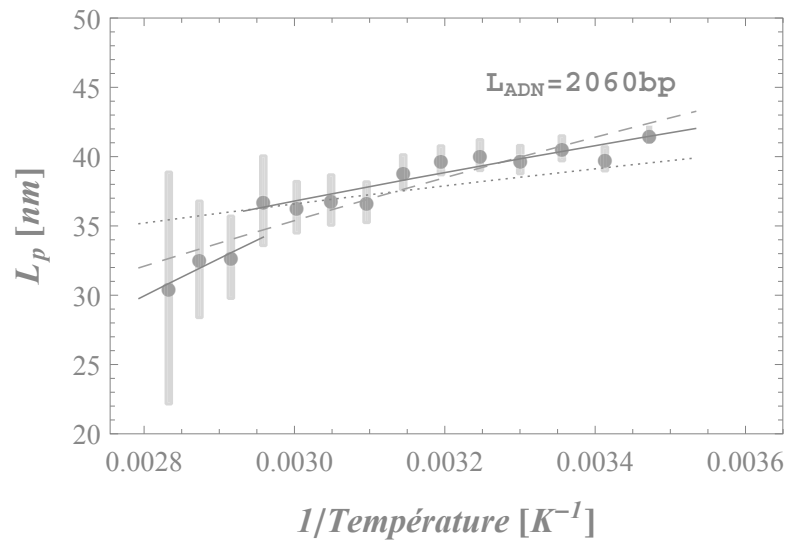


Fig. 7.28 – Variation de  $L_p$  en fonction de  $1/T$ , pour  $L_{ADN} = 2060$  bp. Droite en pointillés noir, ajustement avec un paramètre libre, droite rouge, ajustement avec deux paramètres libres, le tout sur la gamme entière de températures. Droites noires, ajustement sur les gammes restreintes

L'analyse de ces données montre une nette dépendance de  $L_p$  en fonction de la température, contrairement aux études précédentes sur les autres longueurs de molécules d'ADN.

De plus, la rupture de pente, d'un facteur 2, suggère qu'aux alentours de  $65^\circ\text{C}$ ,  $70^\circ\text{C}$ , apparaît une modification de la longueur de persistance apparente de la molécule d'ADN de 2060 bp.

		Expression du fit : $L_p = \frac{\kappa_c}{k_B T} + L_p^c$	
		$\kappa_c$ , en J.nm	$L_p^c$ , en nm
$L_{ADN} = 2060$ bp	$T \in [15 : 80]^\circ\text{C}$	$3.1 \times 10^{-19}$	—
	$T \in [15 : 80]^\circ\text{C}$	$7.3 \times 10^{-19}$	-51.6
	$T \in [15 : 65]^\circ\text{C}$	$4.8 \times 10^{-19}$	-21.0
	$T \in [70 : 80]^\circ\text{C}$	$11.8 \times 10^{-19}$	-105.5

Cette modification est susceptible d'être induite par la formation d'une bulle de dénaturation partielle dans la molécule d'ADN utilisée. Afin de valider ou invalider cette hypothèse nous nous intéressons au profil de température de dénaturation de cette molécule calculé par l'algorithme *MeltSim*.

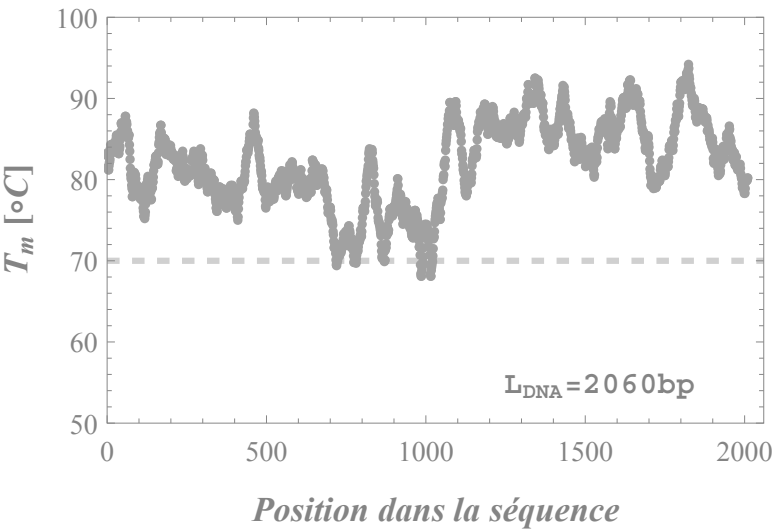


Fig. 7.29 – Profil de la température de dénaturation pour  $L_{ADN} = 2060$  bp en fonction de la position dans la séquence, calculé sur une fenêtre glissante de 50 bp. La droite grise horizontale est un indicateur afin de visualiser la température  $70^\circ\text{C}$

Le profil de température de dénaturation indique que certaines zones de cette molécule d'ADN possèdent une température de dénaturation proche de  $70^\circ\text{C}$ . La variation du paramètre de flexibilité de l'ADN aux alentours de cette température semble donc être provoquée par la dénaturation partielle, sur des séquences courtes, de la molécule de 2060 bp. Par ailleurs, un ajustement des données sur l'ensemble de la gamme de température  $T \in [15 : 80]^\circ\text{C}$  selon la loi stricte Eq. 7.11, c'est à dire un seul paramètre libre selon l'ajustement  $L_p = \frac{\kappa_c}{k_B T}$ , reste très satisfaisant. Ceci laisse supposer que  $\kappa_c$  est indépendant de la température sur l'intervalle considéré.

### 7.6.3 Estimation de la taille de bulle de dénaturation

Notre étude ne nous permet pas d'identifier quantitativement le nombre de bulles de dénaturation ni l'étendue de la zone dénaturée.

En revanche, grâce au modèle *Stiffer-WLC*, nous pouvons estimer qualitativement la taille d'une bulle de dénaturation qui peut être associée à la variation de la longueur de persistance observée. Ce calcul est réalisé dans l'approximation où une seule bulle locale induit cette variation de  $L_p$ . De plus, la position dans la séquence,  $l_{bulle}$  estimée grâce aux positions des minimum de  $T_m$  du profil précédant, Fig. 7.29,  $l_{bulle}$  est donc une position approchée.

Nous supposons que le phénomène de dénaturation se produit, seulement, à la position assimilée aux minima de  $T_m$  proches de  $l = 985$  bp, celui-ci correspond aux deux pics de gauche, quasiment superposés, du profil de température.

La variation de  $L_p = 36.9$  nm pour  $T = 65^\circ\text{C}$ , à  $L_p = 32.9$  nm pour  $T = 70^\circ\text{C}$ , est associée à une diminution de  $R_{ADN}$  de  $R_{ADN_{65^\circ\text{C}}} = 221$  nm à  $R_{ADN_{70^\circ\text{C}}} = 209$ , soit une décroissance de 12 nm, ou  $\sim 6\%$ .

En utilisant le modèle *Stiffer-WLC*, la présence d'une bulle de dénaturation unique à la position  $l = 985$  et dont la longueur de persistance est assimilée à  $L_p = 2L_{p,ss}$ , induisant cette variation de  $R_{ADN}$  conduit à une estimation de taille de bulle de  $L_{bulle} = 135$  bp.

Rappelons que le profil de température indique qu'aux alentours de  $T = 70^\circ\text{C}$ , il apparaît une succession de petites bulles de dénaturation, probablement cinq, et non une seule comme nous l'avions supposé pour le calcul. Ceci contribue donc à une erreur dans notre estimation. Le profil de  $T_m$  laisse supposer qu'il existe une succession de  $N_{bulle}$  bulles de dénaturation locales. Ces dernières, assimilables à une succession de défauts locaux, sont espacées le long de la séquence d'ADN, jusqu'à 100 bp entre elles. Chaque bulle est théoriquement assimilée à une taille respective  $L_{bulle,i}$ . La présence de  $N_{bulle}$  bulles de dénaturation espacées le long de la chaîne du polymère induit une variation de  $R_{ADN}$  plus important que la présence d'une bulle unique de taille  $L_{bulle} = \sum_{i=1}^{N_{bulle}} L_i$ . Il en découle donc que notre estimation sur la taille de la bulle est majorée par ces erreurs. Le nombre réel total de paires de bases dénaturées est donc sûrement inférieur à  $L_{bulle} = 135$  bp. Ce modèle analytique peut être étendu afin d'exprimer la dépendance de  $R_{ADN}$  en présence de plusieurs bulles de dénaturation, ceci n'a pas encore été réalisé.

Forts de cette première étude nous pouvons songer à construire une nouvelle molécule d'ADN de 2000 bp possédant un insert de 60 bp de *Poly(dA)* ainsi que son ADN témoin, comme expliqué précédemment.

## 7.7 COMPARAISON AVEC DES DONNÉES TPM DE LA LITTÉRATURE

Un très récent article de Driessen *et al.* datant de 2014 [DSL<sup>+</sup>14], explore l'influence de la température sur la longueur apparente de l'ADN. Ces résultats leur permettent de remonter à l'extraction de la variation de la longueur de persistance de la molécule d'ADN en fonction de  $T$ .

### 7.7.1 Procédure et données TPM extraites de Driessen *et al.*

Dans cette étude, une molécule d'ADN d'une longueur égale à  $L_{ADN}^D = 685$  bp possède l'une de ses extrémités connectée à une nanoparticule de latex de 460 nm de diamètre,  $R_{par}^D = 230$  nm.

L'impact de l'effet séquence est exploré en considérant différentes constitutions de pourcentage de bases **C-G** dans la séquence de l'ADN  $L_{ADN}^D = 685$  bp. Pour cela les pourcentages suivants 32, 53, et 70% total de bases **C-G** contenus dans l'ADN, sont générés. La solution tampon est constituée de 10 mM HEPES (pH 7.5) et 100 mM NaCl, proche des conditions salines physiologiques.

La gamme de température allant de 23°C à 52°C, par pas de 7°C, est explorée. La distance d'excursion de chaque complexe ADN/particule individuel, notée  $rms$  dans leur article, est mesurée au cours du temps. Pour chaque point de température la durée d'acquisition d'une trajectoire est de 40 s, enregistrée via une caméra possédant une fréquence de 25 Hz et un temps d'exposition de  $T_{ex}^D = 20$  ms.

Leur procédure expérimentale TPM est, en de nombreux points, comparable à la nôtre, hors acquisition à haut-débit comme notre dispositif HT-TPM nous le permet. De plus, leur temps d'exposition  $T_{ex}^D$  est inférieur, de moitié, au nôtre. Leur mesure du paramètre  $rms$  est, bien sûr, directement comparable à notre mesure brute de la distance bout-à-bout des complexes.

Tout d'abord, dans cette étude, plus les molécules d'ADN sont riches en bases C-G plus la valeur de l'amplitude du mouvement augmente, Fig. 7.30 encadré A. De plus, la valeur  $rms$  diminue en fonction de l'augmentation de la température.

Les auteurs extraient  $L_p^D$  par la résolution du problème inverse, étalonnée grâce à leur modèle de simulation de dynamique moléculaire. La variation de la longueur de persistance, extraite de leurs données TPM, diminue proportionnellement à l'augmentation de la température, Fig. 7.30 encadré B. Néanmoins, la diminution de  $L_p^D$  est fortement plus marquée que la diminution théoriquement attendue. En effet, ceci est fortement mis en lumière par le tracé de la dépendance théorique de  $L_p$  en fonction de  $T$  selon l'équation Eq. 7.11, ligne pointillée dans la Fig. 7.30 encadré B.

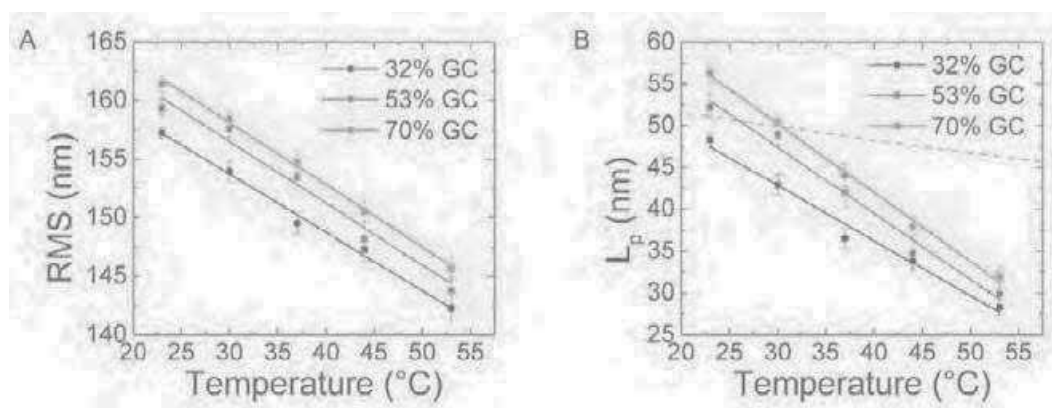


Fig. 7.30 – A Variation de la distance bout-à-bout des complexes ADN/particule en fonction de la température. B Variation de la longueur de persistance en fonction de la température. L'ensemble de ces données correspond à l'étude d'une molécule d'ADN  $L_{ADN}^D = 685$  bp avec différents pourcentages de bases C-G dans la séquence, données TPM de Driessen et al. directement extraites de [DSL<sup>+</sup>14]

Au vu de cette observation, nous avons mis en parallèle leur profil de dépendance de  $L_p$  en fonction de  $T$  et le profil que nous avons obtenu. Pour cela, nous nous focalisons seulement sur les molécules d'ADN contenant 32% et 70% en bases C-G des données extraites de Driessen et al., Fig. 7.31. Ceci nous amène à nous questionner sur les différentes amplitudes de variation observées entre ces deux profils. Il en découle une réflexion sur la nature possible de cette forte dépendance qui est associée aux données TPM de Driessen et al., plus éloignées de la dépendance théorique de  $L_p$  en fonction de  $T$  que la nôtre.

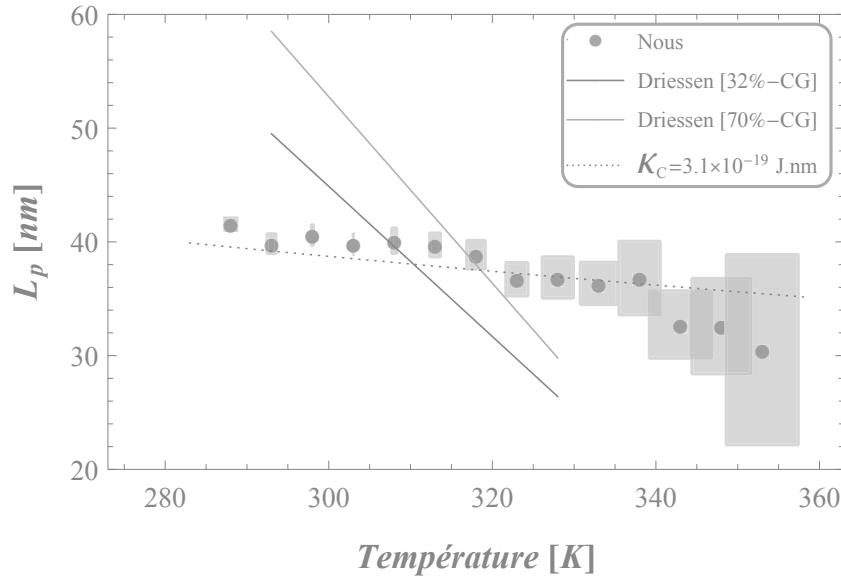


Fig. 7.31 – Variation de la longueur de persistance en fonction de la température. Nos données TPM pour l'ADN 2060 bp sont représentées par les symboles rouge. Ajustement des données TPM extraites de Driessen et al. par l'expression  $L_p^D = L_p^{0,D} - CT$  et dont les valeurs des paramètres de l'ajustement sont directement extraites de [DSL<sup>+</sup>14]. Les ajustements pour l'ADN contenant 32% et 70% en bases C-G sont respectivement représentés par une droite noire et rouge. La courbe en pointillés correspond à l'expression théorique  $L_p = \kappa_c / k_B T$  avec  $\kappa_c = 3.1 \times 10^{-19}$  J.nm, ajustement de nos données pour l'ADN de 2060 avec un seul paramètre libre.

### 7.7.2 Évaluation de l'influence de la viscosité du milieu dans Driessen et al.

Driessen et al. utilisent des simulations numériques basées sur la résolution de l'équation de Langevin, afin d'évaluer l'influence de la viscosité du milieu,  $\eta$ , sur la mesure de  $rms$ .

Cette prospection les amène à conclure que les fluctuations de la viscosité du milieu en fonction de la température n'influent pas sur la mesure des valeurs  $rms$  lors de l'augmentation de  $T$ . Ceci les conduit à n'appliquer aucune correction, dépendante de  $\eta$ , sur leurs données.

Or, comme nous l'avons souligné l'influence de la variation de  $\eta$  en fonction de  $T$  impacte fortement sur le temps de relaxation des complexes ADN/particule. Leur conclusion n'est pas en opposition avec cela. Seulement, dans cet article ils n'ont pas investigué l'effet de  $\eta$  sur  $\tau_{||}$ . Ils ne procèdent à aucune correction de leurs valeurs expérimentales brutes afin de corriger l'effet de moyennage vidéo.

Tout ceci nous amène à penser que leurs données brutes,  $rms$ , nécessitent d'être corrigées de l'effet de moyennage vidéo dû à leur dispositif expérimental. Cette correction rendra directement compte de la variation de  $\eta$  en fonction de l'augmentation de la température.

Afin de valider ou invalider notre hypothèse nous réalisons une correction et une estimation *rapide* de leurs données. En effet, nous ne sommes pas en possession

de leurs données exactes ni des temps de relaxation de leurs systèmes, c'est à cet égard que nous qualifions notre démarche de *rapide*.

### 7.7.3 Estimation du temps de relaxation des données de Driessen *et al.*

La première étape de cette analyse correspond à extraire une courbe de calibration du temps de relaxation des complexes ADN/particule. Via les données actuelles,  $[MTB^{+10}]$ , obtenues grâce à notre modèle de simulation MC-D, nous pouvons estimer la dépendance de  $\tau_{||}$  en fonction de  $L_{ADN}$ , Fig. 7.32.

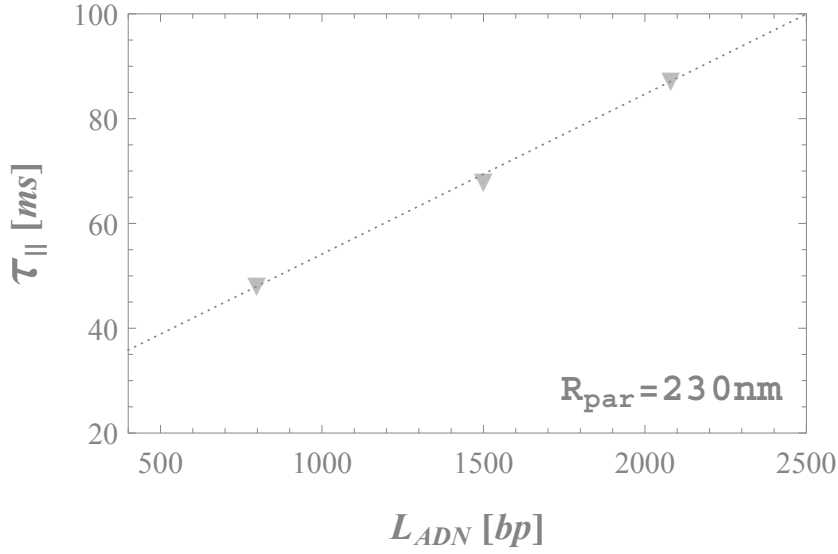


Fig. 7.32 – Variation du temps de relaxation des complexes ADN/particule en fonction de  $L_{ADN}$ , d'après les données publiées dans l'article  $[MTB^{+10}]$ , estimé pour  $R_{par} = 230$  nm, triangles oranges. La droite pointillée correspond à la régression linéaire de ces données.

L'ajustement de ces données par la fonction linéaire  $\tau_{||} = A \times L_{ADN} + B$ , conduit à l'expression suivante :

$$\tau_{||} = 30.5 \times 10^{-3} L_{ADN} + 23.6 \quad (7.12)$$

Pour  $L_{ADN} = 685$  bp et  $T = T_0 = 300^\circ\text{K}$ , l'expression de  $\tau_{||}^D$  ( $D$  pour Driessen *et al.*) est estimée via l'équation Eq. 7.12 comme étant égale à :

$$\tau_{||,685 \text{ bp}}^D = 44.5 \text{ ms} = \tau_{||}^D(T_0) \quad (7.13)$$

### 7.7.4 Estimation des données de Driessen *et al.* corrigées de l'effet lié à $\tau_{||}$

Là encore, nous nous focalisons seulement sur les données associées aux molécules d'ADN contenant 32% et 70% en bases **C-G**.

La première étape de cette analyse consiste à estimer la variation de  $\tau_{||}^D(T)$ , associée à chaque mesure *rms* brute de l'article Driessen *et al.*. Pour cela nous exploitons la dépendance de  $\tau_{||}$  en fonction de  $D_{iff}$  et donc de  $\eta$ . Nous utilisons donc l'équation simple suivante en supposant que  $R_{par} = 230$  nm :

$$\tau_{||}^D(T) = \tau_{||}(T_0) \times \frac{D_{iff}(T_0)}{D_{iff}(T)} = \tau_{||}(T_0) \times \frac{T_0}{T} \frac{\eta(T)}{\eta(T_0)} \quad (7.14)$$

Ainsi, chaque valeur de  $rms$  pourra être associée à une valeur de  $\tau_{||}^D$  via la relation suivante, Eq. 7.15.

$$\tau_{||}^D(T) = 11.37 \times \frac{T}{\eta(T)} \quad (7.15)$$

Enfin, nous appliquons la correction liée à l'effet de moyennage vidéo, dépendant du **temps de relaxation**  $\tau_{||}^D$  des complexes ADN/particule, selon l'équation suivante, [MTB<sup>+</sup>10], où  $T_{ex}^D = 20$  ms :

$$rms_{||}^{D(2)} = rms \left[ 2 \frac{\tau_{||}^D}{T_{ex}^D} - 2 \left( \frac{\tau_{||}^D}{T_{ex}^D} \right)^2 \left( 1 - \exp \left[ -\frac{T_{ex}^D}{\tau_{||}^D} \right] \right) \right]^{(1/2)} \quad (7.16)$$

Ainsi, chaque valeur  $rms$  initiale, est corrigée en fonction de  $\tau_{||}^D(T)$ . Nous pourrions alors extraire la valeur de  $Lp_{||}^{D(1)}$  grâce à l'expression donnée dans la publication de Driessen *et al.*, étalonnée selon leur modèle de simulation :

$$rms_{||}^{D(2)} = 233 - \frac{156}{(1 + 0.08 \times L_p^{D(1)})^{0.45}} \quad (7.17)$$

Nous choisissons cependant d'extraire  $Lp_{||}^{D(2)}$  grâce à nos données de simulation basée sur le modèle d'Echantillonnage Exact. Pour cela, nous évaluons la loi de dépendance de  $R_{Exp||}$  en fonction de  $L_p$  obtenue grâce aux simulations EE, 5.2.3 pour  $L_{ADN} = 685$  bp en imposant la taille des monomères afin de respecter  $a = 6$  bps ce qui nous permet de traduire le diamètre *réel* de l'ADN B, soit  $N = 56 + 1$  monomères. La gamme de longueurs de persistance prospectée est  $L_p \in [107 : 207]$  bp, ce qui correspond à  $L_p \in [30 : 70]$  nm.

	Expression du fit : $\mathbf{R}_{  ,EE}^{2D} = A.L_p^2 + B.L_p + C$			
Cte	$L_{ADN} = 583$ nm	$L_{ADN} = 685$ nm	$L_{ADN} = 1201$ nm	$L_{ADN} = 2060$ nm
A	$-6.00 \times 10^{-3}$	$-7.36 \times 10^{-3}$	$-9.02 \times 10^{-3}$	$-10.70 \times 10^{-3}$
B	1.11	1.35	1.96	2.71
C	110.9	143.1	113.0	162.1

Nous extrayons la variation de la longueur de persistance,  $Lp_{||}^{D(2)}$ , pour l'ADNs de 685 bp utilisée dans Driessen *et al.* grâce à la fonction précédente.

### 7.7.5 Estimations de $L_p^D$ , extraites des données corrigées du moyennage vidéo

Notre prospection *rapide* des données de Driessen *et al.* met en évidence que les corrections liées à l'effet de moyennage vidéo sont la première étape, cruciale, à réaliser avant toute analyse des données TPM. En effet, cet article de 2014 met en évidence une décroissance de  $L_p$  en fonction de la température. Or, contrairement à ce qui est annoncé, cet effet ne peut être justement associée à une variation de la



longueur de persistance, mais correspond à la variation de la viscosité du milieu lorsque  $T$  augmente. En effet, les variations de  $\eta$  en fonction de la température induisent une décroissance du temps de relaxation des complexes ADN/particule qu'il est nécessaire de prendre en considération.

L'ensemble des données de notre proposition d'analyse est contenu dans le tableau suivant :

		Température				
		23°C	30°C	37°C	44°C	53°C
	$\tau_{  }^D$ , bp	49.4	41.3	34.9	29.9	25.4
32% C-G	$rms$ , nm	157.2	154.5	149.5	147.5	142
	$rms_{  }^{D(2)}$ , nm	167.8	167.0	163.7	163.0	161.1
	$Lp_{  }^{D(1)}$ , nm	74.3	71.9	63.4	63.8	57.5
	$Lp_{  }^{D(2)}$ , nm	20.6	19.8	16.8	16.9	14.5
70% C-G	$rms$ , nm	161.4	158.0	155.0	151.0	146.0
	$rms_{  }^{D(2)}$ , nm	172.3	170.7	169.8	167.8	166.2
	$Lp_{  }^{D(1)}$ , nm	89.2	83.7	80.5	74.3	69.9
	$Lp_{  }^{D(2)}$ , nm	25.0	23.4	22.5	20.6	19.1

Un fois ces corrections correctement effectuées, l'influence de la température sur la longueur de persistance,  $Lp_{||}^{D(2)}$ , ne présente plus cette aussi forte diminution, en désaccord avec la théorie.

En effet, notre analyse *rapide*, sans l'exacte connaissance des données ni la prise en compte de leur barre d'erreur, met en exergue que la longueur de persistance varie modérément en fonction de l'augmentation de  $T$ , Fig. 7.33, comme attendu.

Les valeurs de  $L_p$  obtenues sont plus basses qu'attendues, ce qui restera à comprendre, mais la tendance est en assez bon accord avec la variation de  $L_p$  que nous avons observée pour l'ADN de 1201 bp, Fig. 7.33, où aucun phénomène de dénaturation partielle n'est, théoriquement, intervenu sur la gamme de température explorée, comme nous l'avons détaillé précédemment. Ces données sont aussi en accord avec les données de l'ADN de 2060 bp sur la gamme de températures restreintes, sans prendre en contre le domaine de températures où nous avons identifié la formation de bulle de dénaturation.

Pour extraire correctement les valeurs de  $L_p$  associées aux données TPM de Driessen *et al.* il est nécessaire d'appliquer une analyse fine des données brutes afin d'extraire correctement le temps de relaxation des complexes ADN/particule et ainsi, appliquer rigoureusement les corrections associées au moyennage vidéo. Seulement ensuite, l'exploitation des données permet de remonter à la juste variation de  $L_p$  en fonction de la température, non affectée par les variations de la viscosité.

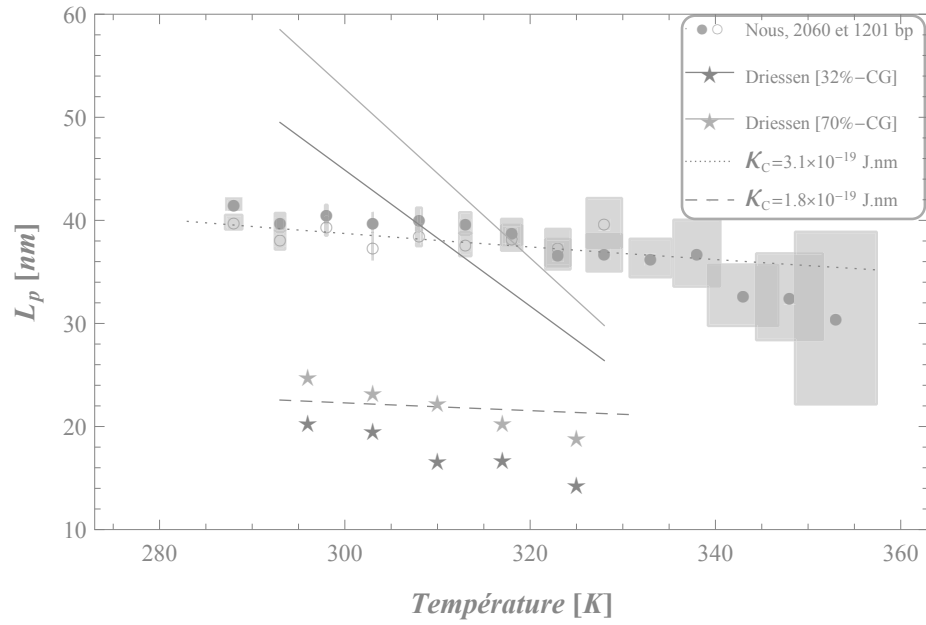


Fig. 7.33 – Variation de la longueur de persistance en fonction de la température. Nos données TPM pour l'ADN 2060 bp sont représentées par des ronds rouges pleins, les données pour l'ADN 1201 bp par des ronds rouges vides. Ajustements des données TPM extraites de Driessen et al. par l'expression  $L_p^D = L_p^{0,D} - CT$  pour l'ADN contenant 32% et 70% en bases **C-G** sont respectivement représentés par une droite noire et rouge. Les données extraites de Driessen et al. et corrigées selon la procédure d'analyse présentée ci-dessus correspondent aux symboles étoilés dont les ADN contenant 32% et 70% en bases **C-G** sont respectivement représentés en noir et en rouge. La courbe en pointillés correspond à l'expression théorique  $L_p = \kappa_c / k_B T$  avec  $\kappa_c = 3.1 \times 10^{-19}$  J.nm et la courbe en tirets noirs avec  $\kappa_c = 1.8 \times 10^{-19}$

## CONCLUSION DU CHAPITRE

Les deux grands processus biologiques clés et vitaux à la transcription et à la réplication de l'information génétique sont dépendants d'une étape essentielle qui implique l'ouverture de la double hélice de l'ADN. Le phénomène de dénaturation induit une perte des multiples interactions faibles maintenant la structure du double brin d'ADN. Ce phénomène peut être provoqué par contraintes extérieures, qui peuvent être liées à la fixation de protéines, par des contraintes thermiques associées à une augmentation de la température, ou bien encore par des contraintes ioniques.

Le phénomène de dénaturation est depuis longtemps étudié afin, entre autre, de mieux comprendre et appréhender les effets qui se produisent à l'intérieur du noyau cellulaire et lors des processus biologiques. La caractérisation fine et, de surcroît, à l'échelle de la molécule individuelle des variations des propriétés physiques et mécaniques du polymère, reste donc un enjeu actuel.

Le développement de modèles basés sur la physique statistique des polymères nous a permis de prospecter et estimer les effets attendus lors d'une dénaturation ponctuelle et locale d'une séquence courte d'ADN. Ceci nous a aidé à définir les limites et les possibilités associées à l'étude de l'influence d'une bulle de dénaturation dans la molécule d'ADN, via notre approche TPM à l'échelle de la molécule unique.

Fort de cette première approche théorique, nous avons pu élaborer une stratégie expérimentale simple. Cette dernière a pour objet l'exploration de l'effet de la température sur une molécule d'ADN. Ceci prend en compte aussi bien l'influence de la présence d'une bulle de dénaturation sur la statistique du système, que l'influence de la température, en amont de ce phénomène de dénaturation, sur la longueur de persistance bio-polymère.

Avant toute chose nous avons mis en place, sur le plan technique, une procédure expérimentale fiable qui assure l'augmentation de température nécessaire à l'évaluation des effets associés. En effet, nous avons dû apporter des solutions pratiques qui nous permettent de maintenir le système à de fortes températures, sans pour autant le modifier physiquement ni altérer notre résolution spatiale et temporelle.

Tout d'abord, nous avons sondé l'impact de la température sur la viscosité du milieu et ainsi étudié son influence sur les propriétés diffusionnelles de nos systèmes ADN/particule. Cet effet souligne l'importance des variations dynamiques associées à la modification de la température du milieu. Confrontés dans un premier temps à notre limite de résolution temporelle, nous avons prospecté différentes tailles de molécules d'ADN afin d'évaluer les bornes associées à cette restriction pour explorer et étudier pleinement ce phénomène.

Ceci nous a permis de mettre en exergue des variations singulières apparaissant sur le temps de relaxation des systèmes ADN/particule mesuré par HT-TPM. Ceci correspond à la signature de phénomènes influant sur des dynamiques des systèmes, que se soit le dépassement de la limite de résolution de notre dispositif expérimental mais surtout la signature de modifications des propriétés physiques de la molécule d'ADN. En effet, de brusques variations du paramètre correspondant à la multiplication du temps de relaxation des systèmes par le coefficient de diffusion peuvent être associées aux phénomènes de dénaturation partielle de la molécule d'ADN. Ce paramètre reste un outil intéressant et une observable complémentaire de la longueur de persistance afin de souligner les fluctuations intervenants sur les propriétés physiques des complexes ADN/particule.

La comparaison avec des données TPM de la littérature a renforcé et souligné l'importance de l'effet de moyennage vidéo sur les mesures expérimentales. En effet, la forte variation de la longueur de persistance obtenue dans l'article de Driessen *et al.* semblent être due, essentiellement, à une diminution de la viscosité du milieu en fonction de l'augmentation de température. En corrigeant l'effet de moyennage vidéo par la prise en compte la variation de  $\tau_{||}$  en fonction de la température, nous avons mis en évidence une variation modérée de  $L_p$  en fonction de  $T$ , plus proche de la dépendance théorique attendue, ainsi que celle observée dans notre étude HT-TPM.

De plus, cette étude nous a permis de mettre en lumière l'influence de l'augmentation de température sur les propriétés physiques de la molécule d'ADN et notamment sur la longueur de persistance. En effet, la diminution de  $L_p$  avec l'augmentation de la température, hors phénomène de dénaturation local et partiel de la molécule d'ADN, suit, comme attendu, une dépendance linéaire inverse. Nous avons pu souligner cette dépendance de  $L_p$  induite par l'effet global correspondant à la variation de  $T$ , sur les deux tailles de molécules d'ADN utilisables dans notre dispositif HT-TPM actuel.

En plus de cela, nous avons pu mesurer et observer de brusques chutes de la longueur de persistance pour de fortes valeurs de températures. Une étude fine du profil de températures de dénaturation obtenu grâce à l'algorithme *MeltSim*, met en évidence l'ouverture locale du double brin d'ADN. La chute observée à  $T = 70^\circ\text{C}$  pour l'ADN de 2060 bp est manifestement associée à la présence de dénaturation partielle au centre de cette molécule.

## BIBLIOGRAPHIE

- [ABLK03] G. Altan-Bonnet, A. Libchaber, and O. Krichevsky. Bubble Dynamics in Double-Stranded DNA. *Physical Review Letters*, 90(13) :138101, April 2003.
- [BBB<sup>+</sup>99] R. D. Blake, J. W. Bizzaro, J. D. Blake, G. R. Day, S. G. Delcourt, J. Knowles, K. A. Marx, and J. SantaLucia. Statistical mechanical simulation of polymeric DNA melting with MELTSIM. *Bioinformatics*, 15(5) :370–375, May 1999.
- [BCPR99] M Barbi, S Cocco, M Peyrard, and S Ruffo. A twist opening model for DNA. *Journal of biological physics*, 24(2-4) :97–114, June 1999.
- [DDPM13] A. K. Dasanna, N. Destainville, J. Palmeri, and M. Manghi. Slow closure of denaturation bubbles in DNA : twist matters. *Phys Rev E Stat Nonlin Soft Matter Phys.*, 2013.
- [DPB93] T. Dauxois, M. Peyrard, and A. Bishop. Entropy-driven DNA denaturation. *Physical Review E*, 47(1) :R44–R47, January 1993.
- [DSL<sup>+</sup>14] R. P C Driessen, G. Sitters, N. Laurens, G. F Moolenaar, G. J L Wuite, and N. Goosen. Effect of Temperature on the Intrinsic Flexibility of DNA and Its Interaction with Architectural Proteins. *Biochemistry*, 53 :6430–6438, 2014.
- [DT62] H. DeVoe and I. Tinoco. The stability of helical polynucleotides : Base contributions. *Journal of Molecular Biology*, 4(6) :500–517, June 1962.
- [Got83] O. Gotoh. Prediction of melting profiles and local Helix stability for sequenced DNA. *Advances in Biophysics*, 16, 1983.
- [GP84] Y. Gao and E. W. Prohofsky. A modified self-consistent phonon theory of hydrogen bond melting. *The Journal of Chemical Physics*, 80(5) :2242, March 1984.
- [Hag88] P J Hagerman. Flexibility of DNA. *Annual review of biophysics and biophysical chemistry*, 17 :265–286, 1988.
- [Hay14] W. M. Haynes. *CRC Handbook of Chemistry and Physics*, 95th Edition. 2014.
- [LFC<sup>+</sup>15] A. Lesne, N. Foray, G. Cathala, T. Forné, H. Wong, and J-M. Victor. Chromatin fiber allostery and the epigenetic code. *Journal of Physics : Condensed Matter*, 27(6) :064114, February 2015.
- [MBC00] B Maier, D Bensimon, and V Croquette. Replication by a single DNA polymerase of a stretched single-stranded DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22) :12002–7, October 2000.
- [MFMC14] F Montel, C Faivre-Moskalenko, and M Castelnovo. Dynamical DNA accessibility induced by chromatin remodeling and protein binding. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 90(5-1) :052717, November 2014.

- [MLCE82] J. C. Marini, S. D. Levene, D. M. Crothers, and P. T. Englund. Bent helical structure in kinetoplast DNA. *Proceedings of the National Academy of Sciences*, 79(24) :7664–7668, December 1982.
- [MM01] A. Montanari and M. Mézard. Hairpin Formation and Elongation of Biomolecules. *Physical Review Letters*, 86(10) :2178–2181, March 2001.
- [MTB<sup>+</sup>10] Manoel Manghi, Catherine Tardin, Julien Baglio, Philippe Rousseau, Laurence Salomé, and Nicolas Destainville. Probing DNA conformational changes with high temporal resolution by tethered particle motion. *Physical biology*, 7(4) :046003, January 2010.
- [PB89] M. Peyrard and A. R. Bishop. Statistical mechanics of a nonlinear model for DNA denaturation. *Physical Review Letters*, 62(23) :2755–2758, June 1989.
- [PDCR09] P. L. Privalov, A. I. Dragan, and C. Crane-Robinson. The cost of DNA bending. *Trends in biochemical sciences*, 34(9) :464–70, September 2009.
- [PS66] D. Poland and H. A. Scheraga. Occurrence of a phase transition in nucleic acid models. *The Journal of chemical physics*, 45(5) :1464–9, September 1966.
- [RYMN96] P. A. Rice, S. Yang, K. Mizuuchi, and H. A. Nash. Crystal structure of an IHF-DNA complex : a protein-induced DNA U-turn. *Cell*, 87(7) :1295–306, December 1996.
- [San98] J. SantaLucia. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences*, 95(4) :1460–1465, February 1998.
- [SCB96] S. B. Smith, Y. Cui, and C. Bustamante. Overstretching B-DNA : The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules. *Science*, 271(5250) :795–799, February 1996.
- [Str85] K. Struhl. Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proceedings of the National Academy of Sciences*, 82(24) :8419–8423, December 1985.
- [Tho93] R. Thomas. The denaturation of DNA. *Gene*, 135(1-2) :77–79, December 1993.
- [TPSW97] B. Tinland, A. Pluen, J. Sturm, and G. Weill. Persistence Length of Single-Stranded DNA. *Macromolecules*, 30(19) :5763–5765, September 1997.
- [WC84] H.-M. Wu and D. M. Crothers. The locus of sequence-directed and protein-induced DNA bending. *Nature*, 308(5959) :509–513, April 1984.
- [ZB59] B. H. Zimm and J. K. Bragg. Theory of the Phase Transition between Helix and Random Coil in Polypeptide Chains. *The Journal of Chemical Physics*, 31(2) :526, July 1959.

# CONCLUSION GÉNÉRALE & PERSPECTIVES

## CONCLUSION

Au cours de ce projet de thèse nous avons exploré avec les outils de la physique les effets de différents phénomènes intervenant sur l'ADN lors de nombreux processus biologiques.

En effet, les ions jouent un rôle majeur sur le déroulement de nombreux processus biologiques affectant la molécule d'ADN. Ces derniers peuvent influencer sur l'activité de liaison de protéines à l'ADN ou encore l'encapsulation de l'ADN dans les capsides virales ou le noyau. L'activité de protéines sur l'ADN est, de surcroît, fréquemment liée à une courbure locale de l'axe de la double hélice. Cette courbure locale peut être présente en raison d'une séquence locale intrinsèquement courbée ou via la capacité des protéines à courber la séquence de la molécule d'ADN sur laquelle elles viennent se fixer. Être capable de caractériser, quantifier et comprendre l'effet des ions présents en solution, la présence d'une courbure locale ainsi que la dénaturation ponctuelle et partielle de la molécule d'ADN sur les conformations de cette dernière est donc crucial et essentiel afin d'approfondir notre compréhension de processus biologiques.

De nombreux travaux tant expérimentaux que théoriques, ont déjà été menés sur ces questions mais celles-ci sont encore largement débattues. En effet, pour y répondre pleinement, doivent notamment être développés des méthodes expérimentales qui ne perturbent pas significativement les conformations de la molécule d'ADN ou encore, le complexe ADN-protéine, ainsi que des modèles théoriques associés qui permettent une analyse précise des données expérimentales ainsi que leur compréhension physique.

L'objectif de ce travail est de développer et proposer des outils expérimentaux et théoriques permettant de décrire l'influence de défauts locaux présents sur la molécule d'ADN, ainsi que l'impact des paramètres physico-chimiques de la solution environnante sur les propriétés physiques et mécaniques de ce bio-polymère.

A cette fin, des données expérimentales ont été acquises à l'échelle de la molécule unique en utilisant pleinement les capacités de la technique à haut-débit de "Tethered Particle Motion" (HT-TPM). La force des approches à l'échelle de la molécule unique réside dans l'accès à l'ensemble de la distribution de comportement d'une molécule individuelle et l'accès direct à la cinétique des processus sans requérir de synchronisation.

La technique TPM utilisée, consiste à enregistrer, au cours du temps, les fluctuations, dues au mouvement brownien, des positions d'une particule accrochée à l'extrémité d'une molécule d'ADN qui elle-même est immobilisée par son autre extrémité sur un support en verre. L'utilisation d'une biopuce permettant la parallélisation des complexes ADN/particule et l'acquisition "à haut débit" de données TPM nous a permis d'obtenir une grande accumulation de statistiques individuelles. Une procédure d'analyse robuste et efficace a été élaborée afin de déterminer les amplitudes de mouvement des assemblages ADN-particules valides.

De plus, touchant aux limites des modèles analytiques classiques de la physique des polymères, nous utilisons et adaptons notre modèle numérique de physique statistique décrivant le bio-polymère ADN à l'échelle mésoscopique. Ainsi, en parallèle nous avons effectué des simulations dans lesquelles la molécule d'ADN est modélisée telle une chaîne constituée de  $N$  monomères de rayons variables, afin de rendre compte des géométries expérimentales associées à l'ADN et à la présence de la particule. Les déplacements du système sont régis par la diffusion brownienne ainsi que par les énergies potentielles d'interactions prenant notamment en compte le module de courbure du polymère ADN. Ceci permet de rendre compte de corrélations des positions le long de la chaîne sur des distances de l'ordre de la longueur de persistance, paramètre clé qui régit la flexibilité et la rigidité intrinsèques du polymère. De plus, ce modèle assimile les contraintes et géométries associées au protocole expérimental. Il nous permet de prospecter et d'extraire la statistique de conformation de l'ADN lors de l'application de variations de ses propriétés physiques, notamment de sonder l'influence de modifications de la longueur de persistance de ce dernier.

Le premier volet de notre étude a porté sur l'effet de la force ionique  $I$  du solvant sur la longueur de persistance  $L_p$ . Ce projet correspond à la prospection d'un effet global, agissant tout le long de la chaîne du polymère, qui influence le paramètre traduisant la rigidité intrinsèque de la molécule d'ADN. L'influence sur  $L_p$  des contre-ions présents en solution est un phénomène qui soulève de nombreuses discussions et interrogations depuis plusieurs décennies. En effet, depuis les années 80 un grand nombre d'études expérimentales ainsi que théoriques se sont succédées. Aujourd'hui, le bilan de ces études fait apparaître une grande diversité de résultats expérimentaux. En effet, ces études mettent en exergue l'existence de différents comportements de l'ADN, notamment aux fortes forces ioniques. Il est de ce fait difficile de dégager un comportement général qui rend compte de ces variations. Les études théoriques les plus célèbres, que sont la théorie d'Odjik-Skolnick-Fixman (OSF) ou encore le modèle de condensation de Manning, ne permettent pas de rendre entièrement compte des variations de la longueur de persistance observées. En effet, ces dernières décrivent convenablement le profil des données expérimentales sur des gammes restreintes de force ionique mais n'expliquent pas le comportement général sur l'ensemble de la gamme de  $I$ .

Ce sujet est donc encore largement ouvert et nous avons décidé d'exploiter les capacités du HT-TPM pour explorer une gamme de forces ioniques comprenant un très grand nombre de mesures aux  $I$  faibles et intermédiaires, et s'étendant jusqu'aux fortes forces ioniques.



Aussi, nous avons prospecté l'impact de la nature des contre-ions en présence, par l'emploi d'ions monovalents  $Na^+$  ainsi que d'ions divalents  $Mg^{2+}$ . Pour consolider notre étude, nous avons sondé ce phénomène sur différentes tailles de molécules d'ADN, allant de 575 bp à 3529 bp afin d'explorer largement la réponse du régime semi-flexible du polymère.

Nous avons pu évaluer et analyser les variations de  $L_p$  extraites des données de HT-TPM, qui font apparaître une décroissance de la longueur de persistance de 55 à 30 nm corrélée à l'augmentation de la force ionique. Une plus forte décroissance est observée pour les ions divalents  $Mg^{2+}$  que pour les ions monovalents  $Na^+$ .

Nous avons effectué une comparaison rigoureuse des valeurs de  $L_p$  à des concentrations salines en ions monovalents proches de celle des conditions salines physiologiques. La variation de  $L_p$  précédemment obtenue dans la littérature sur cette gamme de  $I$  restreinte, soit  $I \in [100 : 200]$  mmol/L, ce qui correspond à  $1/I \in [4 : 12]$  L/mmol, est de l'ordre de  $48 \pm 6$  nm. Cette valeur est en accord avec notre mesure moyenne de  $46.6 \pm 4$  nm, obtenue pour l'ensemble des 6 longueurs d'ADN prospectées durant l'étude.

La comparaison rigoureuse de la variation de  $L_p$  en présence d'ions monovalents  $Na^+$ , nous a amené à souligner les différents comportements de l'ADN en fonction de la valeur de  $I$ . En effet, l'influence des ions monovalents sur les propriétés physiques de l'ADN est, dans la littérature scientifique, sujette à une grande investigation depuis les années 80. Les résultats expérimentaux peuvent être divisés en deux grands groupes définis selon le profil de  $L_p$  à forte force ionique. Lorsque  $I$  augmente dans la gamme des fortes concentrations en ions, le premier groupe observe une lente variation de  $L_p$ . En opposition au second groupe qui lui, observe une variation très significative de  $L_p$  à fort  $I$ . Nos données nous placent en bon accord avec l'augmentation significative de la longueur de persistance sur la gamme des fortes forces ioniques. En revanche, pour la gamme de faibles forces ioniques la variation de  $L_p$  que nous observons est inférieure à celle obtenue par les études expérimentales du second groupe, mais nos données sont en accord avec la variation observée, sur cette gamme, par le premier groupe.

Les théories les plus connues, OSF ainsi que la condensation de Manning, ne permettent pas de rendre compte de la variation de  $L_p$  sur l'ensemble de la gamme de forces ioniques explorées, ni même pour l'ensemble de la nature des ions utilisés. En effet, la théorie OSF évalue qualitativement les variations de  $L_p$  à fort  $I$ , mais cette dernière ne peut aucunement être utilisée pour décrire les fluctuations observées aux  $I$  faibles et intermédiaires. En revanche, sur cette dernière gamme restreinte, la théorie OSF améliorée par la théorie de condensation de Manning décrit de façon plus convenable les fluctuations de  $L_p$  mais seulement en présence d'ions monovalents. L'évaluation des données en présence d'ions divalents s'éloigne, par contre, fortement de la valeur théorique. Là encore, cette théorie ne permet pas la description de  $L_p$  pour toutes les natures des ions présents en solution, ni même sur toute la gamme de sel explorée. Ces deux théories ne sont donc toujours pas suffisantes pour estimer et évaluer correctement la dépendance de  $L_p$  sur une gamme étendue de forces ioniques.

L'approche de Netz et Orland en champ-moyen propose une description adéquate en présence d'ions  $\text{Mg}^{2+}$ , mais n'exprime pas la chute aux fortes concentrations en sel observée en présence d'ions monovalents.

En 2006, Manning reprend sa théorie, et l'implémente d'une approche proposée par Netz, afin de rendre compte de la dépendance de  $L_p$  en fonction de la force ionique sur l'ensemble des concentrations. Cette nouvelle théorie prend en compte la tension électrostatique interne à la molécule due à la répulsion entre les charges portées le long de la chaîne d'ADN qui tend à étirer cette dernière. Les valeurs de  $L_p$  déterminées sur l'ensemble de la plage étendue de forces ioniques ont permis de valider l'approche théorique proposée par Manning en 2006 dans la cas  $\text{Na}^+$ . En effet, cette théorie permet de mieux réconcilier les variations des propriétés physiques de l'ADN à fort et à faible sel. En revanche, cette théorie ne rend pas justement compte de la variation observée en présence d'ions divalents.

Le deuxième volet de cette étude s'intéresse à la quantification de l'angle de courbure locale induite par une séquence spécifique ou la liaison d'une protéine sur la molécule d'ADN. En effet, de nombreuses études ont pu démontrer que la fixation de protéines sur l'ADN peut se traduire par la formation de courbures locales. De même, la présence de séquences spécifiques a son importance et intervient lors de l'initiation et/ou lors du déroulement de nombreux processus biologiques. Les courbures locales induisent des contraintes mécaniques sur la structure de l'ADN. Ceci implique des variations et/ou modifications structurales qui influent sur sa conformation à plus ou moins grande échelle. Ce projet prospecte donc l'impact d'un effet local associé à la présence d'un angle de courbure. La présence de ce dernier a une répercussion sur des paramètres globaux, et peut être observée via la mesure moyenne de la longueur apparente de la molécule d'ADN.

Explorer les effets dus à la présence de courbures locales est encore aujourd'hui, un domaine de recherche. À cette fin, nous nous sommes intéressés à la caractérisation d'une **courbure locale intrinsèque**, utilisée comme **courbure modèle**. Ceci nous a donc permis d'élaborer une méthode de quantification d'un angle de courbure locale et quelconque, induit par une séquence spécifique ou la liaison d'une protéine sur la molécule d'ADN.

L'échantillon modèle a été obtenu en insérant de une à sept séquences CAAAAACGG en phase. Une description théorique de la chaîne d'ADN appelée "kinked Worm-Like Chain" a été proposée afin de rendre compte de l'influence d'un défaut local sur les paramètres statistiques de la chaîne du polymère. Ce modèle K-WLC conduit à une formule simple de la distance bout-à-bout de l'ADN et permet d'extraire la valeur de l'angle de courbure à partir des mesures de HT-TPM. Ainsi, la combinaison des mesure HT-TPM et du modèle K-WLC, a permis de montrer que la séquence CAAAAACGG induit un angle de  $19^\circ \pm 4^\circ$ , en accord avec les données de la littérature.

Ce projet nous a permis de mettre en exergue l'impact de la séquence sur le comportement de la molécule d'ADN. De plus, nous avons pu souligner que l'image que nous pouvions avoir de la molécule d'ADN comme double hélice essentiellement régulière n'est pas convenable. En effet, les variations de  $R_{ADN}$  obtenues en fonction du nombre de séquences courbantes, par rapport à la dépen-

dance théorique attendue, nous a conduit à proposer que toute molécule d'ADN possède une courbure spontanée globale non nulle. Cette approche rend compte d'un meilleur ajustement des mesures HT-TPM par la description théorique. La valeur associée à la courbure globale de la molécule d'ADN de 575 bp que nous avons, ici, utilisée comme référence est donc estimée à  $25^\circ \pm 19^\circ$ .

De plus, nous avons construit une méthode d'analyse simple basée sur le modèle K-WLC de type *clé en main*, afin d'estimer l'angle induit par une courbure locale et ponctuelle pouvant être associé à la présence d'autres objets biologiques. En effet, cette dernière repose sur l'utilisation de seulement 2 mesures ; une mesure de la distance bout-à-bout de l'ADN seul et une mesure de l'ADN incluant la séquence spécifique à étudier et/ou après fixation de protéines. Cette méthode réside sur une préparation et une élaboration de l'expérience plus simple que notre approche précédente qui consistait à construire une série de molécules d'ADN contenant un nombre croissant de séquences spécifiques en phase. La validité et la bonne précision de cette formule est obtenue par l'estimation de l'angle de courbure induit par la séquence CAAAAAACGG utilisée. Cette dernière méthode d'analyse simple conduit à une estimation de l'angle de courbure de  $15^\circ \pm 2^\circ$ , comparable à notre première estimation de  $19^\circ \pm 4^\circ$  obtenue grâce à la mise en relief apportée par l'analyse globale de l'ensemble de la série de molécules courbées.

Nous avons appliqué cette analyse simple sur les données expérimentales obtenues par mesure TPM de la littérature. Ainsi nous avons pu obtenir une estimation de l'angle de courbure induit par la fixation de la protéine IHF proche de  $180^\circ$ , en bon accord avec les données cristallographiques égales à  $\simeq 160^\circ$ . Ceci confirme l'utilité et la cohérence de cet outil, qui permet de quantifier l'effet induit par la présence d'une séquence spécifique et/ou la fixation d'une protéine et d'estimer l'angle de courbure locale et ponctuelle associé.

Enfin, le dernier et troisième volet de ce projet de thèse concerne l'influence de la température sur la rigidité apparente et globale de la molécule d'ADN, jusqu'à la formation d'une zone de dénaturation partielle dans le polymère. Dans cette partie, des résultats préliminaires, mais encourageants, sont présentés. Insistons sur le fait que l'influence de la température sur la molécule d'ADN correspond à la fois, à un phénomène global qui induit une variation des propriétés physiques du polymère, et à un phénomène local qui décrit l'impact de la présence d'une bulle de dénaturation de la molécule. Cette dénaturation partielle, correspondant à la formation d'un défaut local, modifie abruptement les propriétés physiques et intrinsèques de la molécule d'ADN, ce qui se répercute sur la mesure globale de sa longueur apparente ou de sa longueur de persistance. L'influence de la température ainsi que la dénaturation, phénomène intervenant dans les étapes de transcription et de réplication de l'ADN, soulève encore aujourd'hui de nombreuses interrogations. Ces interrogations portent notamment sur la pleine compréhension et quantification de la dynamique de cet effet ainsi que des fluctuations des paramètres statistiques décrivant le polymère.

Afin d'apporter des éléments de réponse nous avons investigué une gamme de températures étendue allant de  $15^\circ\text{C}$  à  $80^\circ\text{C}$ , sondée par pas de  $5^\circ\text{C}$ . Cette étude préliminaire nous a permis de mettre en exergue l'importance des effets de la viscosité du milieu,  $\eta$ , dans ces expériences. En effet, la modification de la tempé-

rature induit une variation de la viscosité du milieu. Or, le temps de relaxation de nos complexes ADN/particules,  $\tau_{||}$ , est proportionnel à  $\eta$ . Ainsi, l'augmentation de  $T$  induit une décroissance de  $\tau_{||}$  qui est une directe conséquence de la décroissance de  $\eta$ .

Cette première observation nous a aidé à définir les limites techniques et expérimentales associées à cette étude. Tout d'abord ceci nous a permis de souligner les limites associées à notre résolution temporelle. Si  $\tau_{||} < 13$  ms, de fortes erreurs sont attendu de la correction de  $R_{TPM}$  de l'effet de moyennage vidéo. Dans cette limite, les données ainsi corrigées de  $R_{Exp||}$  ne sont pas exploitables. Les variations de la viscosité du milieu en fonction de la température imposée nous mènent à franchir cette limite sans avoir pu prospecter l'ensemble de la gamme de températures voulue. En effet, pour l'ADN de 583 bp, possédant un temps de relaxation de  $\sim 20$  ms à  $T = 20^\circ\text{C}$ , la limite associée à notre résolution temporelle est franchie dès la température de  $T = 30^\circ\text{C}$ , impliquant toute analyse fine.

De plus, nous avons pu définir la taille limite de molécules d'ADN qui nous permet de prospecter l'ensemble de la gamme de températures voulue, en tenant compte des contraintes induites par notre limite de résolution temporelle. Nous dispositif expérimental nous impose d'utiliser une molécule d'ADN de longueur de contour minimale de l'ordre de 2000 bp. Par exemple, pour l'ADN de 1201 bp, possédant un temps de relaxation de  $\sim 35$  ms à  $T = 20^\circ\text{C}$ , la limite associée à notre résolution temporelle est franchie à la température de  $T = 60^\circ\text{C}$ . Pour l'ADN de 2060 bp, cette limite n'est pas franchie à  $T = 80^\circ\text{C}$ .

Forts de la mise en évidence des contraintes associées à notre limite de résolution temporelle ainsi que de la longueur de la molécule d'ADN minimale nécessaire, nous avons pu prospecter entièrement la gamme de températures voulue et quantifier les variations des propriétés physiques intrinsèques d'une molécule d'ADN de 2060 bp. Ainsi, nous avons pu mettre en exergue l'influence de la température sur la longueur apparente ainsi que sur la longueur de persistance de la molécule d'ADN, allant même jusqu'à identifier la signature de la formation d'une dénaturation partielle dans cette molécule. En effet, il apparaît une chute de la longueur de persistance aux alentours de  $T = 65^\circ\text{C}$ . Cette rupture de pente doit correspondre à la formation de petites bulles de dénaturation dans la molécule d'ADN comme le profil de température de dénaturation obtenu via l'algorithme *MeltSim* le suggère. L'utilisation du modèle S-WLC, nous permet d'associer cette décroissance à une variation locale de la rigidité du polymère. Ainsi, nous avons pu estimer et assimiler cette chute de  $L_p$  à la formation d'une bulle de dénaturation ponctuelle de 135 bp. La décroissance de  $L_p$  est proportionnelle à l'inverse de la température que nous avons obtenue, est en accord avec la théorie, contrairement aux données de Driessen *et al.* qui présentent une forte décroissance. En effet, les variations de la longueur de persistance obtenues dans cet article semblent être dues à la combinaison de deux effets distincts : la variation de  $L_p$  en fonction de  $T$  mais aussi, et surtout, la variation de  $\tau_{||}$  en fonction de  $T$ . Il est donc bien nécessaire de procéder à une correction de l'effet de moyennage vidéo des données TPM afin d'isoler l'influence de  $T$  sur la longueur de persistance. La décroissance de  $L_p$ , ainsi obtenue, est alors cohérente avec la théorie.

De plus, tout cela nous a permis d'exploiter les informations fournies par le calcul de la fonction  $F_{TPM}$ , correspondant à la multiplication de la viscosité du milieu par le temps de relaxation de nos complexes ADN/particule. Cette fonction  $F_{TPM}$  semble être un véritable outil qui peut permettre de mettre en lumière une variation singulière intervenant sur  $\tau_{||}$ . Dans un premier temps ceci nous a permis de souligner les limites associées à notre résolution temporelle. En effet, lorsque que la limite de résolution temporelle est atteinte, la dépendance de  $F_{TPM}$  est inversée. Dans second temps, cette inversion de pente place  $F_{TPM}$  tel un outil privilégié qui nous permet de sonder et visualiser un phénomène de dépendance singulière du temps de corrélation. En effet, la dénaturation partielle de la molécule d'ADN de 2060 bp est associée à une variation de la dépendance de  $L_p$  en fonction de  $T$  aux alentours de  $T = 65^\circ\text{C}$ . À cette température, la fonction  $F_{TPM}$  présente aussi une particularité, l'inversion de pente de cette dernière semble être une signature de l'apparition d'un phénomène singulier.

En conclusion, la technique HT-TPM et l'accumulation des données obtenues par l'exploitation de la bio-puce, nous ont permis d'observer et de quantifier, avec une bonne précision, la dépendance de  $L_p$  en fonction de  $T$  et la formation de dénaturation partielle dans la molécule d'ADN. Avec l'ensemble des études précédentes, nous avons pu observer l'influence et extraire la valeur de l'angle de courbure induit par la présence d'une séquence courbante courte et spécifique ainsi que quantifier finement l'influence de la concentration en sel, pour différentes natures d'ions présents en solution, sur les propriétés physiques de la molécule d'ADN.

## PERSPECTIVES

D'un point de vue essentiellement technique, des améliorations peuvent être apportées à notre procédure expérimentale HT-TPM. Le premier aspect de ces améliorations est d'optimiser notre capacité d'accumulation des données via notre bio-puce dans toutes les conditions de force ionique et/ou de température. Le second aspect repose sur l'amélioration de notre résolution temporelle pour permettre de sonder, par exemple, la dynamique des bulles de dénaturation formée dans la molécule d'ADN.

Dans un premier temps, l'ensemble des étapes du protocole permettant la fixation spécifique de la molécule d'ADN sur la surface solide ainsi que l'accroche de la particule à l'autre extrémité du bio-polymère peuvent être légèrement modifiées. En effet, l'utilisation de groupements correspondant à des liaisons fortes peut être envisagée que se soit pour la fonctionnalisation de la surface et/ou de la particule. Ainsi, nous pouvons considérer différents groupements pour les extrémités modifiées de l'ADN. Les interactions spécifiques biotine/streptavidine et Dig/Anti-Dig peuvent être remplacées par des groupements Thiol(SH)/Maleimide et/ou Amide(NH<sub>2</sub>)/Carboxylique(COOH). Par exemple, une fonctionnalisation via un groupement COOH permet une liaison covalente avec un groupement NH<sub>2</sub>. Ce type de fonctionnalisation peut être réalisé sur les deux zones de jonction des extrémités de l'ADN. Ceci pourra nous permettre d'augmenter la force des interactions mises en jeu et ainsi diminuer le taux de décrochage observé à faible et forte force ionique, ainsi que lors de la forte variation de la température du milieu.

La forte capacité d'accumulation des données via notre bio-puce *TPM-on-a-chip*, nous permet de pleinement explorer l'ensemble des conformations accessibles

à l'ADN et d'obtenir une forte pertinence en terme de poids statistique, lors de chacune de nos mesures. En revanche, nous sommes limités en termes de dynamique et ce, sur deux plans distincts. Tout d'abord, comme nous l'avons vu lors de la prospection de l'influence de  $T$  sur  $L_p$ , nous sommes limités par notre dispositif expérimental via notre temps d'exposition. Ceci a rendu impossible l'exploration et la mesure de données expérimentales, en terme de statistique, car le temps de relaxation de nos complexes ADN/particule était inférieur à notre limite de résolution temporelle. Ensuite, en dehors de l'accumulation statistique de données, nous sommes limités dans l'exploration de phénomènes possédant une dynamique plus rapide que notre résolution temporelle. Une première solution est donc d'augmenter notre fréquence d'acquisition, ce qui nous permet à la fois d'abaisser notre limite de résolution temporelle et de sonder des dynamiques plus rapide. Seulement, il est nécessaire de conserver un bon rapport signal/bruit. Ceci nous amène à songer à utiliser des nanoparticules ayant une bonne section efficace, l'utilisation de nanoparticules d'or de 80 nm de diamètre est donc toute indiquée. Ces dernières accélèrent d'autant plus la dynamique de notre système qu'elles sont petites. Comme nous l'avons souligné dans ce manuscrit, ceci nous a confronté à la limite directement induite par notre dispositif de détection et de la perte du bon rapport signal/bruit par l'utilisation de la caméra actuelle. Le développement du dispositif expérimental, combinant l'utilisation d'une nouvelle caméra possédant une vitesse d'acquisition plus rapide couplée à l'emploi de nanoparticules d'or de diamètre inférieur, est actuellement en cours d'optimisation à l'IPBS.

Dans ce cadre-là, il nous sera donc sûrement utile de raffiner nos modèles dynamiques de simulations. Ceci nous permettra de prévoir convenablement les résolutions temporelles associées aux phénomènes que nous souhaitons sonder et évaluer la faisabilité des mesures expérimentales. Comme nous l'avons souligné, le modèle **Helical** permet une approche fine de la dynamique et des mécanismes régissant la fermeture d'une bulle de dénaturation. De plus, le modèle MC-D rend compte et modélise de façon très convenable les conformations accessibles au polymère associées à notre dispositif expérimental. La combinaison de nos deux modèles, **MC-D** et **H**, qui n'a pour l'instant pas été réalisée, peut nous permettre une entière quantification de la présence et de la dynamique associée à une bulle de dénaturation tout en respectant la géométrie de notre approche expérimentale. Le développement de ce modèle multi-échelle est à l'étude.

Dans l'analyse même, lors de l'extraction de  $L_p$  de nos mesures expérimentales, quel que soit l'effet prospecté, nous avons pu juger de l'erreur induite par les méthodes de soustraction de la contribution de la bille et d'extraction de  $R_{ADN}$  et de  $L_p$  basées sur les modèles physiques des polymères. Nous avons pu pallier ces erreurs grâce à l'apport de nos modèles de simulations qui modélisent correctement la géométrie expérimentale ainsi que les différentes contraintes liées à l'existence des différents volumes exclus, dus à la présence du mur, de la particule avec la chaîne ou de la chaîne avec elle-même. Améliorer les descriptions théoriques et les équations analytiques associées afin de rendre compte convenablement de ces effets sans induire de nouveaux biais et d'importantes barres d'erreur est donc crucial et nécessaire. En effet, l'apport des simulations a été prouvé, mais transcrire cet apport dans une expression analytique pouvant être utilisée simplement, loin de la mise en œuvre plus lourde associée aux séries de simulations numériques nécessaires par thématique de projets, deviendrait un

outil commun et plus pratique d'utilisation. Ces outils pourraient être utilisés par un large profil de chercheurs venant de l'ensemble des domaines s'intéressant aux bio-polymères, qu'ils viennent du domaine de la biologie moléculaire, de la bio-physique et qu'ils soient théoriciens ou expérimentateurs. Il reste donc un effort considérable à fournir afin de répondre à cette problématique.

Comme nous l'avons vu, encore aujourd'hui, l'influence de la force ionique sur la longueur de persistance n'est pas entièrement comprise. Seule la nouvelle théorie de Manning s'approche d'une analyse quantitative des mesures expérimentales. Par contre, l'impact et de l'importance de la nature des ions n'est pas pleinement prise en compte par les équations théorique dont nous disposons actuellement. De plus amples investigations semblent nécessaires afin de tester l'influence des ions en fonction de leur nature. Nous avons observé une nette différence entre les mesures de  $L_p$  en fonction des ions présents en solution, monovalents ou divalents. Pour tester si cet effet est principalement dû à leur valence, nous pourrions comparer l'effet induit par la présence d'ions monovalents  $K^+$  à celui obtenu en présence d'ions  $Na^+$ . Sur ce même principe la comparaison de la dépendance de  $L_p$  en présence d'ions  $Ca^{2+}$  par rapport à celle obtenue en présence d'ions magnésium serait intéressante. De plus, ceci pourrait nous permettre de mettre en exergue si la taille des ions en présence influe sur les propriétés physiques de la molécule d'ADN, de façon significative ou non. Ce domaine laisse encore de nombreuses questions en suspens et offre de nombreux champs d'investigations avec notre outil expérimental.

Cette étude nous a aussi permis de souligner une forte augmentation du temps de relaxation des complexes ADN/particule aux très forte concentration en sel. Cet effet mesuré sur le paramètre  $\tau_{||}$  est seulement observable en présence d'ions monovalents et uniquement pour deux longueurs de molécule d'ADN. Ce phénomène reste pour l'instant non expliqué. Des mesures plus étendues sur la gamme de forte concentration en sel en présence d'ions monovalents, ainsi que d'ions divalents, pourraient nous permettre de sonder plus précisément cet effet. En complément des mesures de rayons hydrodynamiques sur la gamme de fortes forces ioniques prospectées peuvent être envisagées. En effet, les mesures que nous avons actuellement réalisées sur la gamme de  $I$  allant de 10.1 mmol/L à 510.1 mmol/L, rendent compte d'un rayon hydrodynamique des complexes ADN/particule constant à  $\pm 1$  nm, ce qui ne nous permet pas de conclure sur l'effet souligné par cette brusque variation de  $\tau_{||}$  à très fort  $I$ . De plus, il serait intéressant de considérer l'influence de la viscosité du milieu que de fortes concentrations en sel pourraient modifier.

L'étude sur l'influence de séquences courbantes accumulées en phase selon le pas de l'hélice a mis en lumière l'existence d'une fluctuation de nos données expérimentales autour de l'effet moyen attendu. En effet, cette fluctuation n'est pas correctement prise en compte lors de nos différents ajustements que ce soit en considérant seulement l'influence d'une courbure locale et ponctuelle dans l'ADN ou en considérant la présence de cette même courbure combinée à l'existence d'une courbure globale sur l'ensemble de la molécule d'ADN. Notre première hypothèse afin d'expliquer ce phénomène est que la construction même de la série des ADN utilisés supprime différentes séquences courtes dans la molécule ce qui peut induire des suppressions de faibles courbures de façon non homogène. Afin de prospecter cet effet nous pouvons grâce à notre modèle de simulation par

échantillonnage exact sonder l'effet d'angles de courbure,  $\theta$ , aléatoires mais homogènes tout le long de la molécule d'ADN sur la mesure de la longueur apparente de cette dernière. Ces réalisations numériques dites en *désordre gelé* peuvent nous permettre d'explorer les valeurs et limites de  $\theta$  qui induisent, en moyenne, une molécule d'ADN régulière droite et/ou possédant une courbure globalement non nulle. La deuxième étape consiste donc à imposer une courbure locale et ponctuelle, correspondant à la courbure induite par la séquence spécifique CA<sub>6</sub>CGG, dans ces simulations en *désordre gelé*. L'étude de la dispersion des données pourra nous aider à mieux comprendre et quantifier le phénomène de fluctuation des données que nous avons pu observer.

Par ailleurs, la mise en relief des molécules possédant des séquences A-tract assemblées en phase avec le molécule dont les séquences spécifiques sont construites en opposition de phase semble indiquer une augmentation modérée de la rigidité pour la dernière séquence d'ADN. Dans un premier temps, nous pourrions sonder l'influence de l'accumulation de séquences spécifiques CA<sub>6</sub>CGG construites en opposition de phase. En effet, actuellement nous avons seulement construit une molécule d'ADN constituée de 4 séquences disposées en opposition de phase, nous pourrions l'étendre à des constructions de 6 et 7 séquences. Ceci pourrait nous permettre de sonder et quantifier l'influence de cette construction en opposition de phase sur les rigidités de la molécule d'ADN. Or, comme nous l'avons déjà souligné, notre dispositif expérimental ne nous permet pas de discriminer facilement une courbure locale d'une modification de la rigidité. Afin de mesurer proprement cette variation, nous pourrions sonder, dans un deuxième temps, le régime élastique du polymère. La méthode classique afin de tester la rigidité du polymère est d'appliquer un étirement sur cette dernière, comme notamment les mesures d'AFM. Des expériences d'étirement par flux de la molécule d'ADN pourrait nous permettre de distinguer l'influence de ces deux causes sur les propriétés physiques de la molécule d'ADN. Sous couvert de quelques modifications de notre dispositif expérimental, des expériences d'étirement par flux peuvent être réalisées. La comparaison et la concordance entre ces mesures d'étirement de l'ADN sous force et des mesures HT-TPM actuelles pourrait contribuer à aller plus loin dans l'étude des variations des propriétés physiques de la molécule d'ADN. En effet, la mise en parallèle de ces résultats pourrait nous permettre de sonder différents régimes de réponses élastiques de l'ADN et, peut-être discriminer plus facilement entre une courbure locale et une modification de la rigidité.

Dans l'ensemble de ces deux parties nous avons aussi pu mettre en exergue l'influence de la séquence même de l'ADN sur ses propriétés physiques. En effet, la présence d'une séquence *Poly(dA)* de 60 bp influe fortement sur les propriétés du polymère soumis à une variation de force ionique, et ce quelle que soit la nature des ions présents en solution. De plus, la présence d'une séquence *Poly(dA)* peut être assimilée à une variation locale de la courbure de l'ADN ou encore à une modification locale du paramètre de rigidité du polymère. Une telle modification, qui peut-être considérée comme un défaut local dans la structure du double brin d'ADN, peut influencer le comportement global de la molécule. En effet, ici, une modification de 60 bp, soit un dixième de la séquence, influe directement sur la mesure des propriétés physiques de l'ensemble de la molécule d'ADN de 583 bp. Or, à ce jour nous n'avons pas pleinement compris les rouages de la nature même de cet effet. Là encore nous ne pouvons discriminer facilement entre l'existence d'une courbure ponctuelle ou d'une variation locale de la rigidité, ceci nécessite



de plus amples travaux. L'observation de cet effet séquence interroge sur les propriétés et les répercussions que peut engendrer une modification structurale locale de l'ADN. Là encore, la combinaison de notre méthode expérimentale avec d'autres approches semble intéressante.

L'étude préliminaire sur l'effet de la température et la présence de dénaturation partielle dans la molécule d'ADN est encourageante et laisse un large champ d'investigation pour le futur via notre approche HT-TPM. Il semble donc intéressant de construire une molécule d'ADN d'une taille minimale de 2000 bp, construite afin d'offrir la possibilité d'une dénaturation partielle et contenue dans la séquence, donc non proche des extrémités de la molécule. Une telle molécule permettrait la formation d'une bulle de dénaturation locale et ponctuelle autour de la température de dénaturation, qui doit, idéalement, être distincte de la température de dénaturation de toute autre zone de la séquence. Ceci pourrait nous permettre de pleinement explorer l'influence de la température et l'influence de la formation d'une zone dénaturée sur le comportement global de la molécule d'ADN. Dans de telles conditions, il pourrait aussi être très intéressant de réaliser, en plus de la montée en température, une redescende de température. En effet, ceci pourrait permettre de vérifier si les mesures réalisées en  $T$  décroissant sont cohérentes avec celles réalisées en  $T$  croissant. Tout phénomène d'hystérésis serait donc mis en lumière. Dans cette optique, il sera donc nécessaire de ne pas atteindre la température critique de dénaturation de l'ensemble de la molécule, rendant toute re-hybridation de la molécule d'ADN délicate.

Pour finir, l'ensemble des effets prospectés séparément peuvent être combinés. En effet, ceci nous permettrait de sonder si la présence de séquences courbantes spécifiques influe sur la dépendance de  $L_p$  en fonction de la force ionique. Les constructions en phase et en opposition de phase n'ont peut-être pas le même comportement en fonction de  $I$  ou de la nature des ions. Ceci peut amener des éléments de réponse quant aux zones d'interaction des ions et contre-ions présents en solution avec la molécule d'ADN. De plus, la variation de la longueur de persistance de ces ADN spécifiques en fonction de la température peut être différente.

L'ensemble de ce travail nous a permis de prospecter une large étendue d'effets, aussi bien globaux que locaux, impactant sur les propriétés physiques et mécaniques de la molécule d'ADN. L'ensemble de ces effets induisent de fines variations des paramètres physiques intrinsèques de la molécule d'ADN et nécessite donc une forte accumulation de données. De plus, l'ensemble de ces résultats ouvrent sur de nombreuses questions qui soulignent l'importance de la séquence même de l'ADN, ce qui fait écho avec les différentes séquences spécifiques qui interviennent *in vivo* lors de nombreux processus biologiques. De plus, jusqu'à présent les physiciens se sont peu intéressés aux effets induit par la séquence même de la molécule d'ADN sur les propriétés physiques de cette dernière. L'ensemble de cette étude nous a permis de montrer que la séquence influe fortement sur le comportement global du polymère. De surcroît, notre méthode HT-TPM est un véritable outil qui nous permet de sonder l'effet séquence, sous différentes conditions. Tout ceci laisse entrevoir le nombre et la richesse des études qu'il sera possible de mener grâce à notre approche expérimentale HT-TPM intimement couplée à nos modèles numériques qui rendent compte correctement de la géométrie conformationnelle des complexes ADN/particule, et nous permettent de sonder la dynamique associée.



## Cinquième partie

### Annexes



## SOMMAIRE

A	PROTOCOLE . . . . .	313
A.1	Détails du protocole expérimental . . . . .	313
A.2	Structure Ordonnées de protéine, moule et timbre de PDMS . . . . .	324
B	ADNs UTILISÉS . . . . .	324
B.1	Effet du sel - ADN 575 bp . . . . .	324
	Paramètres de la synthèse pour $L_{ADN} = 575$ bp . . . . .	324
	Séquence 575 bp . . . . .	324
B.2	Effet du sel - ADN 583 bp . . . . .	324
B.3	Effet du sel - ADN 639 bp . . . . .	325
	Paramètres de la synthèse pour $L_{ADN} = 639$ bp . . . . .	325
	Séquence 639 bp . . . . .	325
B.4	Effet du sel - ADN 1201 bp . . . . .	325
	Paramètres de la synthèse pour $L_{ADN} = 1201$ bp . . . . .	325
	Séquence 1201 bp . . . . .	325
B.5	Effet du sel - ADN 2060 bp . . . . .	326
	Paramètres de la synthèse pour $L_{ADN} = 2060$ bp . . . . .	326
	Séquence 2060 bp . . . . .	326
B.6	Effet du sel - ADN 3259 bp . . . . .	326
	Paramètres de la synthèse pour $L_{ADN} = 3259$ bp . . . . .	326
	Séquence 3259 bp . . . . .	327
B.7	Courbure - ADN 575 bp, sans et avec un nombre croissant de sé- quences <b>A-tract</b> : $CA_6CGG$ . . . . .	328
	Paramètres de la synthèse pour $L_{ADN} = 575$ bp, sans courbure . . . . .	328
	Séquence 575 bp, $6An0$ , sans courbure . . . . .	328
	Paramètres de la synthèse pour $L_{ADN} = 575$ bp, avec courbure . . . . .	329
	Séquence 575 bp, $6Ani^P$ ou $O$ , avec courbure . . . . .	329
B.8	Courbure - ADN 575 bp, sans et avec un nombre croissant de sé- quences <b>A-tract</b> : $TA_3TC$ . . . . .	329
	Paramètres de la synthèse pour $L_{ADN} = 575$ bp, avec et sans courbure . . . . .	329
	Séquence 575 bp, $6An0$ , sans courbure . . . . .	329
B.9	Courbure - ADN 575 bp, sans et avec un nombre croissant de sé- quences <b>A-tract</b> : <i>Rand</i> . . . . .	329
B.10	Dénaturation - ADN 583 bp, avec et sans coeur <i>polyA</i> . . . . .	330
	Paramètres de la synthèse pour $L_{ADN} = 583$ bp . . . . .	330
	Séquence 583 bp <i>PolyA</i> . . . . .	330
	Séquence 583 bp . . . . .	330
C	SCRIPT <i>Mathematica</i> D'ANALYSE DES DONNÉES . . . . .	331
D	SCRIPT DES PROGRAMMES NUMÉRIQUES . . . . .	364
D.1	Logigramme du modèle Monte-Carlo Dynamique (MC-D) . . . . .	364
D.2	Logigramme du modèle Échantillonnage Exacte (EE) . . . . .	365
E	DONNÉES GÉNÉRALES . . . . .	366
E.1	Données TPM, courbe de calibration . . . . .	366
E.2	Données simulation, courbe de calibration . . . . .	366
F	DONNÉES SUR L'EFFET DE LA FORCE IONIQUE . . . . .	367
G	DONNÉES SUR L'EFFET D'UNE COURBURE LOCALE . . . . .	375
H	DONNÉES SUR L'EFFET DE LA TEMPÉRATURE . . . . .	375



## A PROTOCOLE

### A.1 Détails du protocole expérimental

#### Protocole de préparation de lamelles v1

**But :** recouvrir des lamelles de verre de silanes terminés par fonction epoxy. (Pour lier par microcontact printing des protéines par les fonctions  $\text{NH}_2$  libres).

**Matériel :** lamelles de verre : 5 (24x36).

#### 1)-Premier rinçage:

- Rincer rapidement, à la pipette d'éthanol, les 5 lamelles
- Immerger les lamelles dans de l'éthanol pur **et** mettre le tout au bain à ultrasons pdt 5min
- Les sécher (minifuge à lame ou  $\text{N}_2$  ou à l'étuve à  $110^\circ\text{C}$  pendant 5min).

#### 2)-Rinçage au sulfochromique:

- Les immerger dans une solution d'acide sulfochromique (Hotte salle 2) au moins 30 min (jusqu'à 2h).
- Les rincer abondamment à l' $\text{H}_2\text{O}$  mQ (extensivement, 1L). Pour le transport laisser de l'eau au fond du cristalliseur.
- Les sécher (minifuge à lame, ou  $\text{N}_2$ ) **et** 5 min à  $115^\circ\text{C}$ .

#### 3)-Fonctionnalisation avec la fonction epoxy:

- Incuber les lamelles dans une solution d'époxidation pendant 1h30 :
  - 19.5 mL isopropanol (l'utiliser pour laver la verrerie)
  - 0.2 mL  $\text{H}_2\text{O}$  mQ
  - 0.5 mL 3-Glycidoxypropyldiméthoxyméthylsilane GPDS (Sigma 539252)
  - 20  $\mu\text{L}$  N,N-Diméthylbenzylamine (Sigma 185582)
- Les immerger 5min dans isopropanol **et** bain à ultrasons.
- Les nettoyer rapidement à la pissette à EtOH
- Les sécher (minifuge à lame, ou  $\text{N}_2$ ), **et** 15min à  $115^\circ\text{C}$ .

**Stockage :** jusqu'à 1 mois, sous vide.

## DÉTAILS Protocole de préparation de lamelles v1

**But :** recouvrir des lamelles de verre de silanes terminés par fonction epoxy.  
(Pour lier par microcontact printing des protéines par les fonctions  $\text{NH}_2$  libres).

**Matériel :** lamelles de verre : 5 (24x36).

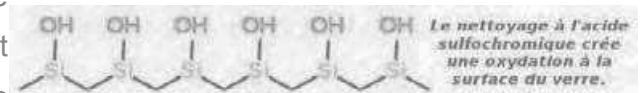


### 1)-Premier rinçage:

- Permet d'enlever les plus grosses poussières et impuretés adsorbées sur la surface.
- Le bain d'éthanol pur est utilisé plusieurs fois, d'où l'utilité d'un pré-rinçage. L'utilisation d'ultra-son permet de décrocher un maximum d'impuretés.

### 2)-Rinçage au sulfochromique:

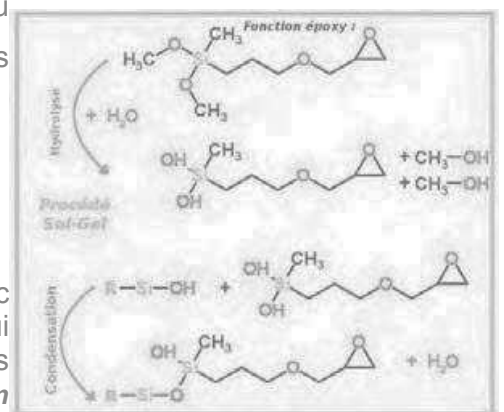
- L'acide sulfochromique crée des oxydations de surface successives qui permettent de piéger et relarguer les impuretés avant le greffage des fonctions époxys sur la surface.



- Le fait de laisser un fond d'eau dans le cristalliseur lors du transport permet d'éviter de recontaminer les surfaces des lamelles.

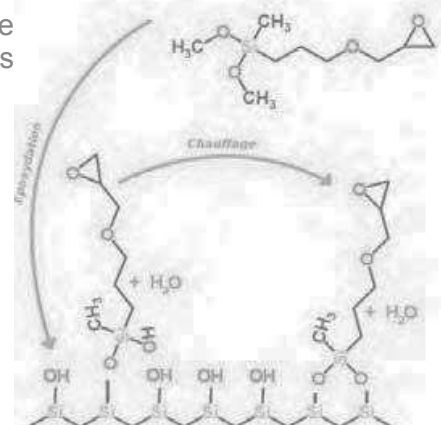
### 3)-Fonctionnalisation avec la fonction epoxy:

- Ici l'époxydation correspond à fonctionnaliser le substrat avec une molécule chimique GPDS possédant la fonction époxy qui permet la liaison covalente avec les groupements amines libres des neutravidines (cf *Protocol de fabrication de timbres en PDMS/C.Microcontact Printing*)



Utilisation de la molécule GTDS (= Tri oxyde méthyle) favorise la création de structure 3D, alors que nous nous recherchons un résultat surfacique plan et homogène.

**Stockage :** jusqu'à 1 mois, sous vide.





## Protocole de préparation de lamelles v2

**But :** recouvrir des lamelles de verre de silanes terminés par fonction epoxy. (Pour lier par microcontact printing des protéines par les fonctions  $\text{NH}_2$  libres).

**Matériel :** lamelles de verre : 5 (24x36).

### **1)-Premier rinçage:**

- Rincer rapidement, à la pipette d'éthanol, les 5 lamelles
- Puis, mettre le tout au bain à ultrasons pdt 5min
- Les sécher (minifuge à lame ou  $\text{N}_2$  ou à l'étuve à  $110^\circ\text{C}$  pendant 5min).

### **2)-Rinçage a l'hellmanex:**

- Immerger les lamelles dans de l'hellmanex 4% **et** mettre le tout au bain  $60^\circ\text{C}$  pendant 5min
- Puis, mettre le tout au bain à ultrasons pdt 5min
- Les sécher (minifuge à lame ou  $\text{N}_2$ ) **et** 5 min à  $115^\circ\text{C}$ .

### **3)-Fonctionnalisation avec la fonction epoxy:**

- Incuber les lamelles dans une solution d'époxidation pendant 1h30 :
  - 19.5 mL isopropanol (l'utiliser pour laver la verrerie)
  - 0.2 mL  $\text{H}_2\text{O}$  mQ
  - 0.5 mL 3-Glycidoxypropyldimethoxymethylsilane GPDS (Sigma 539252)
  - 20  $\mu\text{L}$  N,N-Dimethylbenzylamine (Sigma 185582)
- Les immerger 5min dans isopropanol **et** bain à ultrasons.
- Les nettoyer rapidement à la pissette à EtOH
- Les sécher (minifuge à lame, ou  $\text{N}_2$ ), **et** 15min à  $115^\circ\text{C}$ .

**Stockage :** jusqu'à 1 mois, sous vide.

## Protocole de préparation de lames

But : recouvrir une face des lames de verre de PEG.

Matériel : lames de verre ( ~ 45x25mm , angle coupé toujours en haut à droite).

### 1)-Premier rinçage:

- Rincer rapidement, à la pipette d'éthanol, les 5 lamelles
- Immerger les lamelles dans de l'éthanol pur **et** mettre le tout au bain à ultrasons pdt 5min
- Les sécher (minifuge à lame ou N<sub>2</sub> ou à l'étuve à 110°C pendant 5min).

### 2)-Rinçage au sulfochromique:

- Les immerger dans une solution d'acide sulfochromique (S2) au moins 30 min (jusqu'à 2h).
- Les rincer abondamment à l' H<sub>2</sub>O mQ (extensivement, 1L) **et** laisser dans eau
- Les sécher (minifuge à lame, ou N<sub>2</sub>) **et** 5 min à 115 °C.

### 3)-Fonctionnalisation avec la fonction thiol

- Incuber dans une solution de thiolation pendant 1 nuit.
  - 21.5 mL méthanol
  - 1 mL H<sub>2</sub>O mQ
  - 0.2 mL acide acétique
  - 1mL 3-mercaptopropyltriméthoxy-silane (Sigma, 175617)
- Les immerger 5 min dans méthanol et bain à ultrasons.
- Les sécher (minifuge à lame, ou N<sub>2</sub>), puis 5 min à 115°C.

### 4)-Passivation de la surface avec du PEG

- Déposer ~30 µL (par lame de verre) d'une solution à 10 mg/mL de PEG (Méthoxypolyéthylène glycol maléimide -malPEG 5000, Sigma 63187- dans Hepes - H7523- 10mM pH6.5) sur une feuille de parafilm adhérent à la paillasse.
- Déposer les lames de verre sur le parafilm et laissez incubez 1h.
- Les rincer extensivement à H<sub>2</sub>O mQ
- Les sécher (minifuge à lame, ou N<sub>2</sub>), puis 5 min à 115 °C

Stockage : jusqu'à 1 mois, sous vide, ou jusqu'à 3-4 utilisations dans chambre.

## Protocole de préparation de lames

But : recouvrir une face des lames de verre de PEG.

Matériel : lames de verre ( ~ 45x25mm , angle coupé toujours en haut à droite).

### 1)-Premier rinçage:

- Permet d'enlever les plus grosses poussières et impuretés adsorbées sur la surface.
- Le bain d'éthanol pur est utilisé plusieurs fois, d'où l'utilité d'un pré-rinçage. L'utilisation d'ultrasons permet de décrocher un maximum d'impuretés.

### 2)-Rinçage au sulfochromique:

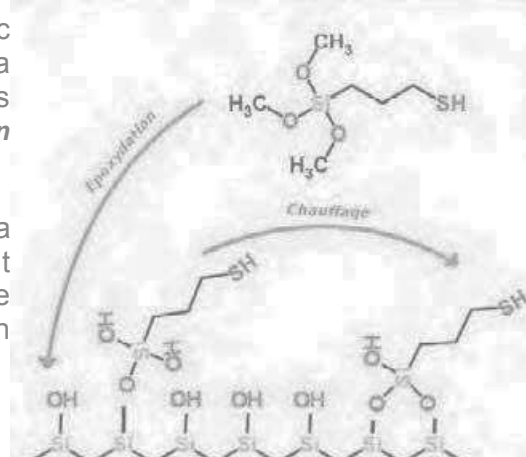
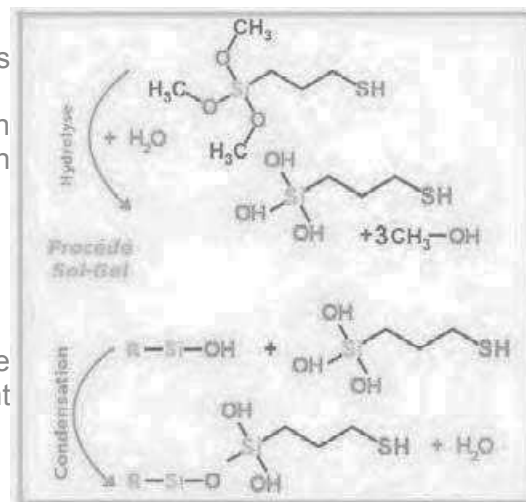
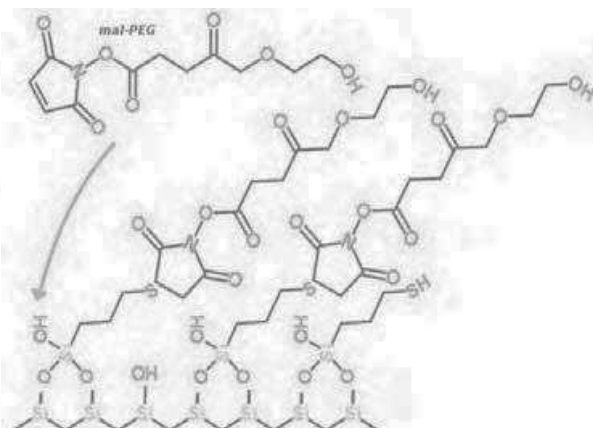
- L'acide sulfochromique crée une oxydation de surface successive qui permet de piéger et relarguer les impuretés avant le greffage des fonctions époxy sur la surface.

### 3)-Fonctionnalisation avec la fonction thiol

- Ici la thiolisation correspond à fonctionnaliser le substrat avec une molécule chimique possédant la fonction SH. qui permet la liaison covalente avec les groupements sames libres des neutravidines (cf *Protocol de fabrication de timbres en PDMS/C.Microcontact Printing*)

Utilisation de la molécule *Tri oxyde méthyle* car elle reste la seule dans le commerce actuellement . Comme précédemment le □□□ doit favoriser la création de structure 3D, là encore l'utilisation de □□devrait permettre de créer une fonctionnaliation de surface plane et homogène.

### 4)-Passivation de la surface avec du PEG



- Le PEG permet de passer la lame supérieure pour limiter les interactions non-spécifiques entre ADN/nPs et la surface de verre. Le PEG utilisé possède un groupement  $\text{CH}_2\text{CH}_2\text{OH}$  qui permet d'interagir avec le groupement  $\text{Si-OH}$  du silane utilisé dans l'étape précédente.

Stockage : jusqu'à 1 mois, sous vide, ou usqu'à 3-4 utilisations dans chambre.

## Fonctionnalisation des nPs LATEX avec AC anticorps Anti-Digoxigénine

But : lier des anti-Digoxigénine (Roche- 11093274910) à des billes de latex.

*Le protocole est adapté de celui d'Ademtech (pour les billes magnétiques). Il fonctionne pour 0.5 mg de billes carboxylate de 300nm de Ø de chez Merck-Estapor.*

### **1)-Re-suspendre les nPs dans buffer:**

- rincer un tube eppendorf avec 100µL d'Activation Buffer 1X (AB1X, Ademtech).
- dans ce tube vidé, ajouter 50µL de billes F1-XC030 (Merck-Estapor, ≈0.5mg).
- centrifuger pendant 9 min à 10000 rpm = 9300g à ~22°C (@RT).
- éliminer le surnageant, resuspendre le culot avec 100 µL d'AB1X, vortexer.
- faire l'étape de centrifugation / resuspension **3** fois.

### **2)-Première activation:**

- ajouter 40 µL d'EDC (1-ethyl-3-(3-diméthylaminopropyl)carbodiimide, Invitrogen E-2247), solution à 4 mg/mL dans AB1X, vortexer.
- incuber pendant 10 min à 37°C.

### **3)-Fonctionnalisation avec anti-corps:**

- ajouter 75 µL d'une solution d'anti-Digoxigénine (200 µg/mL dans un tampon PBS pH7.4 – Euromedex-ET330), vortexer.
  - incuber pendant 2 h à 37°C.
  - ajouter 300 µL d'une solution de BSA (Sigma-B4501-, 0.5 mg/mL dans AB1X), vortexer
  - incuber pendant 30 min à 37°C.
  - centrifuger pendant 15 min à 10000 rpm (@RT).
  - éliminer le surnageant, resuspendre le culot avec 100 µL de Storage Buffer 1X (SB1X, Ademtech), vortexer.
  - centrifuger pendant 8 min à 10000 rpm (@RT).
  - éliminer le surnageant, resuspendre le culot avec 100 µL de SB1X (Ademtech), vortexer.
  - centrifuger pendant 8 min à 10000 rpm (@RT).
  - éliminer le surnageant, resuspendre le culot avec 112 µL de SB1X (Ademtech), vortexer.
- Pour 112 µL, la concentration finale de billes fonctionnalisée est de 500 pM
- Soniquer les billes : eppendorf dans la glace, sonde à sonication (puissance 50, interrupteur 25/50 sur 25), 4 x 5 s espacées de 15 s.

stockage : jusqu'à 2 mois, à 4°C. Avant chaque utilisation, soniquer 5 min dans un bain à sonication.

## Formation des complexes nanoparticules - ADN.

But : lier les ADN double brins digoxigénine aux particules fonctionnalisées.

Tampon A (TpA) :

Nom	Concentration Solution Stock	Concentration solution finale	Volume solution Stock pour 30mL final
PBS	1□□	1□	3□□
Pluronic 127	□□□ g/□□	1□ g/□□	1□□□□
BSA	□□□ g/□□	□1□ g/□□	□□□□
Tween 20		□1□	3□□□
Triton X100		□□□□□	1□□□
H <sub>2</sub> O milliQ			□□□□□

- Rincer extensivement la chambre avec 20-30 volumes de TpA (la chambre a un volume de ~10μL).
- Incuber la chambre avec ~10μL de TpA, pendant 30 min @RT.
  - Préparation des particules fonctionnalisées avec ADN :
  - Il faut que les particules et l'ADN a la même concentration, 50pM, pendant 20min @RT
  - Exemple, pour une chambre : la chambre fait 10μL finale  
Il faut donc 5μL à 100pM de billes avec 5μL d'ADN à 100pM.  
La solution stock de billes est à 500pM, il faut donc 1μL de stock dans 4μL de TpA.  
L'ADN 2080 bdp est aliquotée, 0,5μL à 31nM : il faut donc la diluer avec 155μL de TpA, et en prélever 5μL.
- Rincer extensivement la chambre avec 20-30 volumes de TpA (la chambre a un volume de ~10μL).
- Introduire les 10μL de billes fonctionnalisées ADN dans la chambre
- Laisser incuber à 4°C dans la chambre humide pendant 4h minimum ou O/N.
- Rincer extensivement la chambre avec 20-30 volumes de TpA (la chambre a un volume de ~10μL) pour enlever les billes en suspension avant d'observer.

## Protocole de fabrication de timbres de PDMS

Utilisation des puces de silicium d'AMO GmbH (100 mm<sup>2</sup>). La surface d'intérêt est un carré de 16mm<sup>2</sup> au centre.

Matériau de la puce de silicium.

Procédure de fabrication



Les puces de silicium sont recouvertes d'une pellicule protectrice :

- Les immerger dans un cristalliseur pendant 10 min dans de l'acétone, sous agitation.
- Les rincer à l'isopropanol.
- Les sécher par flux d' $N_2$ , puis 5 min à 115°C.
- Passer les puces refroidies au Plasma (Pico, Diener, Oxygène 0.25 mbar 5 min à 200W).
- Les immerger aussitôt dans une solution d'OTS (Trichloro(octadécyl)silane, Sigma 104817) à 1% dans du toluène, pendant 5 min. Immersion dans un petit cristalliseur (6cm Ø), sous flux constant d' $N_2$ .
- Les rincer d'abord avec un peu de toluène, puis de l'isopropanol.
- Les sécher par flux d' $N_2$ , puis 5 min à 115°C.

Les puces hydrophobes peuvent être utilisés au max. ~ 50 fois.

Nettoyage au HF

Quand des traces de PDMS réticulé restent dessus, il faut les nettoyer au HF.

- Verser dans un bûcher en téflon 30 mL d'acide sulfurique (pur, min. 95%), puis 5 mL d'eau peroxydée ( $H_2O_2$  à 30%, Sigma 216763)  **S**
- Immerger les puces dans cette solution pendant 5 min.
-  Les rincer à l' $H_2O$  mQ, les sécher sous flux d' $N_2$ .
- Les immerger dans un bûcher en téflon contenant une solution de HF (ratio 6/1 de fluorure d'ammonium  $NH_4F$  40%-Sigma 09738- et d'acide fluorhydrique HF 49%-Sigma 47559-) pendant 20 s.

**Matériau de la puce de silicium** (VWR-ALFA44542.15)  **S**

- Les rincer à l' $H_2O$  mQ.
- Les sécher sous  $N_2$  puis 15 min à 115°C

Les puces peuvent de nouveau être fonctionnalisées avec l'OTS.

## Bibliothèque de motifs de S

### 1)-Préparation du PDMS:

- Mélanger **10 unités** de PDMS Sylgard® 184, avec **1 unité** d'agent réticulant (ratio 10/1) dans un becher en téflon (les blanc).
- Mélanger très vigoureusement **et** ensuite dégazé jusqu'à élimination de toutes les bulles visibles.

Stockage : Le PDMS mélangé est stockable à -20°C pendant ~2 mois.

### 1)-Préparation des timbres en PDMS:

- Placer les puces de Si hydrophobes sur un morceau parafilm sur une plaque de verre
- Disposer ~150µL de PDMS sur chaque puce.
- Mettre à cuire (65°C pendant au moins 12h), faire bien attention à ce que la plaque de verre sur laquelle sont les puces est bien horizontale (niveaux à bulle)

Stockage : dans les boites de culture 48 puits, indéfiniment.

## Protocole de contact

But : dépôt de protéine sur lamelle

- Couper la lamelle de verre époxydée en 2 (pointe diamant).
- Nettoyer les surfaces par un flux d' $N_2$ .
- Poser sur la ½ lamelle une feuille de silicone prédécoupée (Clear Silicon Sheet, 70338-29 de EMS, 0.25mm épaisseur, 2 chambres de 3x11 mm).
- Séparer le timbre de PDMS de la puce de silicium (à l'aide de pinces).
- Poser le timbre (face arrondie vers le bas) sur une lame de verre.
- Découper le timbre de manière à obtenir des cubes de 1x1 mm<sup>2</sup> à 2x1 mm<sup>2</sup> ne contenant sur la face supérieure que les motifs.
- Incuber sur le timbre ~15-30 µL de neutravidine (marquée ou non Invitrogen-A6373 ou A2666) à 20 µg/mL dans un tampon PBS (pH 7.4 Euromedex-ET330) pendant 1 min.
- Rincer immédiatement avec H<sub>2</sub>O (à la pissette) et passez sous un flux d' $N_2$  (10sec).
- Retourner le tampon et poser sa surface encrée sur la lamelle pendant 1 min.
- Enlever le tampon en faisant attention à ne pas le traîner sur la surface.
- Placer la lame de verre, les trous aux extrémités de la chambre (angle coupé en haut à droite).

Stockage : jusqu'à 1 mois, à 4°C.

## Protocole de Microcontact Printing (Print des plots)

**But :** dépôt des protéines : **neutrophile** dans la chambre sur les lamelles de verres.

**Matériel :** lamelles de verre : 5 (24x36)  
timbre de PDMS

### **1)-Préparation des lamelles :**

- sur une feuille de parafilm adhérent à la pailleasse couper la lamelle de verre époxydée en 2 grâce à une pointe diamant.
- nettoyer les surface par un flux d' $N_2$  pour enlever les poussières
- poser sur la  $\frac{1}{2}$  lamelle une feuille de silicone prédécoupée (Clear Silicon Sheet, 70338-29 de EMS, 0.25mm épaisseur, 2 chambres de 3x11 mm).

### **2)-Préparation du timbre PDMS :**

- séparer le timbre de PDMS de la puce de silicium (à l'aide de pinces).
- poser le timbre (face arrondie vers le bas) sur une lame de verre.
- découper le timbre de manière à obtenir des cubes de  $1 \times 1 \text{ mm}^2$  à  $2 \times 1 \text{ mm}^2$  ne contenant sur la face supérieure que les motifs (au final mieux vaut moins que trop).
- incuber sur le timbre 20  $\mu\text{L}$  de neutravidine (marquée ou non Invitrogen-A6373 ou A2666) à 20  $\mu\text{g/mL}$  dans un tampon PBS (pH 7.4 Euromedex-ET330) pendant 1 min.
- rincer immédiatement  $H_2O$  (à la pissette) et passez sous un flux d' $N_2$  (10sec).
- retourner le tampon et posez sa surface encrée sur la lamelle pendant 1 min.
- enlever le tampon en faisant attention à ne pas le traîner sur la surface.
- placer la lame de verre, les trous aux extrémités de la chambre (angle coupé en haut à droite).

**stockage :** jusqu'à 1 mois, à 4°C.



## Microcontact Printing (Print des trous) Façon IBM

But : Dépôt d'une protéine sur lamelle, ais avec des tailles de spots ne permettant pas l'utilisation du  $\mu$ contact printing classique.

Prendre l'Ac X Digoxigénine marqué FITC

- Prendre solution stock Ac XDig (20 $\mu$ L à 200 $\mu$ g/mL).
- Le diluer dans le PBS jusqu'à C=100 $\mu$ g/mL.
- Couper un cube de PDMS de  $\approx 3 \times 3 \text{ mm}^2$ .
- Incuber une goutte d'Ac XDig 100 $\mu$ g/mL sur le stamp de PDMS, face contre boîte de pétri vers le haut, pdt 1h.

- Laver extensivement à l'eau milliq et sécher au N<sub>2</sub>.

- Prendre une puce de silicium sans pdms.
- Mettre sur une lame de  $\mu$ scope.
- Activer la puce de silicium au plasma 02 pdt 5 min / 0.25mb / 200 W (100%).
- Couper une lamelle époxy, souffler au N<sub>2</sub> pour enlever d'éventuels résidus et monter sur un parafilm et un masque de silicone (pour guider).

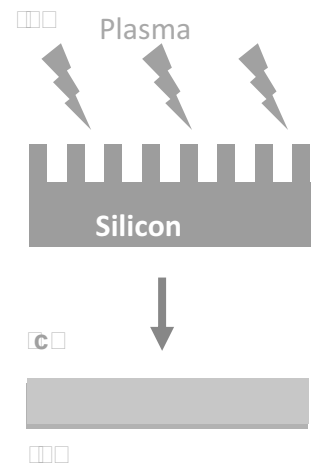
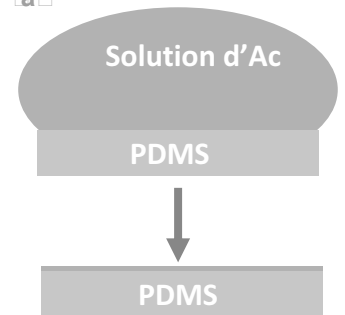
- DES QUE LA PUCE DE SILICIUM SORT DU PLASMA CLEANER, appliquer le PDMS sur le motif de la puce pdt 1h.

- Appliquer le stamp sur la lamelle epoxy pdt 1h.

- Monter la lamelle entre un nouveau masque de silicone et une lame en verre trouée fonctionnalisée PEG.

- Vérifier le stamping en fluo.

a



e

## A.2 Structure Ordonnées de protéine, moule et timbre de PDMS

Représentation des surfaces de silicium lithographiées qui nous servent de moule, ou de parton inversé pour nos timbres de PMDS. Ceci permet d'ordonnée le dépôt des protéine sur la surface en une succession d'îlots séparées.

Le détail de structure que nous possédons sont définit en suivants.

Les taille possibles des îlots sont les suivantes :

- 1060 nm
- 1015 nm
- 980 nm
- 450 nm
- 280 nm

Les taille possibles des aires vides séparant ces îlots sont les suivantes :

- 1.4  $\mu\text{m}$
- 3.4  $\mu\text{m}$ .

## B ADNs UTILISÉS

### B.1 Effet du sel - ADN 575 bp

Paramètres de la synthèse pour  $L_{ADN} = 575$  bp

(A) :  $L_{ADN} = 575$  bp

- Plasmide  $\rightarrow$  pTOC1
- Oligo-1  $\rightarrow$  F575
- Oligo-2  $\rightarrow$  R575

#### Séquence 575 bp

```
ATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTCAATATTATTGAAGCATTT
TCAGGGTTATTGTCTCATGAGCGGATGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTT
AGAAAAATAAACAAATAGGGGTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAA
GAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTCTT
CAAGAATTCTCATGTTTGACAGCTTATCATCGATAAGCTTGGTGCGCATAATGTATATTATGTT
AAATCTGGTGCGCATAATGTATATTATGTTAAATCTGGTGCGCATAATGTATATTATGTTAAAT
GTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGA
CTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGC
GCTCTGGGTCATTTTCGGCGAGGACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTCGCTT
CGGTATTTCGGAATCTTGCACG
```

### B.2 Effet du sel - ADN 583 bp

Se référer à la séquence B.10.

### B.3 Effet du sel - ADN 639 bp

Paramètres de la synthèse pour  $L_{ADN} = 639$  bp

□ :  $L_{ADN} = 639$  bp

- Plasmide → pTOC1
- Oligo-1 → F639
- Oligo-2 → R639

#### Séquence 639 bp

```
AGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGT
CTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAA
AAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCC
CTTTCGTCTTCAAGAATTCTCATGTTTGACAGCTTATCATCGATAAGCTTGGTGCGCATAATGTATATTATGTTA
AATCTGGTGCGCATAATGTATATTATGTTAAATCTGGTGCGCATAATGTATATTATGTTAAATGTCGACCGATGC
CCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACT
GTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCAATTTTCGGCGAGGACCGCTTTCGC
TGGAGCGCGACGATGATCGGCCTGTCGCTTGCAGTATTTCGGAATCTTGCACGCCCTCGCTCAAGCCTTCGTAC
TGGTCCC GCCACCAAACGTTTCGGCGAGAAGCAGGCCATTATCGC
```

### B.4 Effet du sel - ADN 1201 bp

Paramètres de la synthèse pour  $L_{ADN} = 1201$  bp

□ :  $L_{ADN} = 1201$  bp

- Plasmide → pTOC1
- Oligo-1 → F1201
- Oligo-2 → R1201

#### Séquence 1201 bp

```
CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGATGCGGGCAGCGAGTTGCTCTTGCCCGGCGTCAACAC
GGGATAATACCGCGCCACATAGCAGAACTTTAAAAAGTGCTCATCATTGGAAAAACGTTCTTCGGGGCGAAAACTC
TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCAACTGATCTTCAGCATCTTTT
ACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGCGACGAC
GGAAATGTTGAATACTCATACTCTTCCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGG
ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTG
ACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAA
GAATTCTCATGTTTGACAGCTTATCATCGATAAGCTTGGTGCGCATAATGTATATTATGTTAAATCTGGTGCGCAT
AATGTATATTATGTTAAATCTGGTGCGCATAATGTATATTATGTTAAATGTCGACCGATGCCCTTGAGAGCCTTCA
ACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGC
AACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCAATTTTCGGCGAGGACCGCTTTCGCTGGAGCGCGACGATGA
TCGGCCTGTGCTTGCAGTATTTCGGAATCTTGCACGCCCTCGCTCAAGCCTTCGTCACTGGTCCC GCCACCAAAC
GTTTCGGCGAGAAGCAGGCCATTATCGCCGGCATGGCGGCCGACGCGCTGGGCTACGTCTTGCTGGCGTTCGCGA
CGCGAGGCTGGATGGCCTTCCCCATTATGATTCTTCTCGCTTCCGGCGGCATCGGGATGCCCGGCTTGCAGGCCAT
GCTGTCCAGGCAGGTAGATGACGACCATCAGGGACAGCTTCAAGGATCGCTCGCGGCTCTTACCAGCCTAACTTC
GATCATTGGACCGCTGATCGTCACGGCGATTATGCCGCCTCGGCGAGCACATGGAACGGGTTGGCATGGATTGTAG
```

## B.5 Effet du sel - ADN 2060 bp

Paramètres de la synthèse pour  $L_{ADN} = 2060$  bp

ⓓ :  $L_{ADN} = 2060$  bp

- Plasmide → pTOC1
- Oligo-1 → F2060
- Oligo-2 → R2060

Séquence 2060 bp

CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAA  
GGGCCGAGCGCAGAAGTGGTCCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGAAGC  
TAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCACACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTAC  
GCTCGTCGTTTGGTATGGCTTCATTAGCTCCGGTCCCAACGATCAAGGCGAGTTACATGATCCCCATG  
TTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTATC  
ACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTG  
GTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAAC  
ACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA  
AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTT  
CAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGG  
GAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTCAATATTATTGAAGCATTATCAGG  
GTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACA  
TTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTTATTATCATGACATTAACCTATAAAAAATAGGCG  
TATCACGAGGCCCTTCGTCTTCAAGAATTCTCATGTTTGACAGCTTATCATCGATAAGCTTGGTGCACATA  
ATGTATATTATGTTAAATCTGGTGCGCATAATGTATATTATGTTAAATCTGGTGCGCATAATGTATATTATGT  
TAAATGTGCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGTCTCTTCGGTGCGCGCGGGCATGACTA  
TCGTGCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCA  
TTTTCGGCGAGGACCGCTTCGTCTGGAGCGCGACGATGATCGGCCTGTGCTTGGGTTATTCGGAATCTTGC  
ACGCCCTCGCTCAAGCCTTCGTCACTGGTCCCGCCACCAAACGTTTCGGCGAGAAGCAGGCCATTATCGCC  
GGCATGGCGGCCGACGCGCTGGGCTACGTCTTGCTGGCGTTTCGCGACGCGAGGCTGGATGGCCTTCCCCATT  
ATGATTCTTCTCGCTTCCGGCGGCATCGGGATGCCCGCGTTGACAGGCCATGCTGTCCAGGCAGGTAGATGAC  
GACCATCAGGGACAGCTTCAAGGATCGCTCGCGGCTCTTACCAGCCTAACTTCGATCATTGGACCGCTGAT  
CGTCACGGCGATTTATGCCGCCTCGGCGAGCACATGGAACGGGTTGGCATGGATTGTAGGCGCCGCCCTAT  
ACCTTGTCTGCCTCCCCGCGTTGCGTCGCGGTGCATGGAGCCGGGCCACCTCGACCTGAATGGAAGCCGGC  
GGCACCTCGCTAACGGATTACCACTCCAAGAATTGGAGCCAATCAATTCTTGCGGAGAACTGTGAATGCG  
CAAACCAACCCTTGGCAGAACATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACGCGGCGCATCTCGGG  
CAGCGTTGGGTCTTGCCACGGGTGCGCATGATCGTGCTCCTGTGCTTGAGGACCCGGCTAGGCTGGCGGGG  
TTGCCTTACTGGTTAGCAGAATGAATCACCGATACGCGAGCGAACGTGAAGCGACTGCTGCTGCAAAACGT  
CTGCGACCTGAGCAACAACATGAATGGTCTTCGGTTTCCGTGTTTCGTAAAGTCTGGAAACGCGGAAGTCA

## B.6 Effet du sel - ADN 3259 bp

Paramètres de la synthèse pour  $L_{ADN} = 3259$  bp

ⓔ :  $L_{ADN} = 3259$  bp

- Plasmide → pTOC1
- Oligo-1 → F3259
- Oligo-2 → R3259

**Séquence 3259 bp**

GCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC  
TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCT  
TGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAA  
GCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAG  
CGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCC  
TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTC  
ATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA  
ATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTA  
TCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCCGTCGTGTAGATAACTACG  
ATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGG  
CTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTCTGCAA  
CTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTT  
AATAGTTTGCACAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTACGCTCGTCGTTTGGTAT  
GGCTTCATTTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAA  
AAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTGAGAAGTAAGTTGGCCGCAGTGTTATCACT  
CATGGTTATGGCAGCACTGCATAATTCTCTTACTGTTCATGCCATCCGTAAGATGCTTTTCTGTGA  
CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCC  
GGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAAGTGCTCATCATTGAAAA  
ACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCC  
ACTCGTGCACCCAACTGATCTTCAGCATCTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAA  
CAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATA  
CTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTT  
GAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAAGTGCCACCT  
GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCT  
TTCGTCTTCAAGAATTCTCATGTTTGACAGCTTATCATCGATAAGCTTGGTGCGCATAATGTATA  
TTATGTTAAATCTGGTGCGCATAATGTATATTATGTTAAATCTGGTGCGCATAATGTATATTATGT  
TAAATGTGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGC  
ATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGC  
AGCGCTCTGGGTCATTTTCGGCGAGGACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTGCG  
TTGCGGTATTCCGAATCTTGACGCCCCCTCGCTCAAGCCTTCGTCACTGGTCCCGCCACCAAACGT  
TTCGGCGAGAAGCAGGCCATTATCGCCGGCATGGCGGCCGACGCGCTGGGCTACGTCTTGCTGGC  
GTTGCGGACGCGAGGCTGGATGGCCTTCCCCATTATGATTCTTCTCGCTTCCGGCGGCATCGGGAT  
GCCCGCGTTGCAGGCCATGCTGTCCAGGCAGGTAGATGACGACCATCAGGGACAGCTTCAAGGAT  
CGCTCGCGGCTCTTACCAGCCTAACTTCGATCATTGGACCGCTGATCGTCACGGCGATTATGCCG  
CCTCGGCGAGCACATGGAACGGGTTGGCATGGATTGTAGGCGCCGCCCTATACCTTGTCTGCCTCC  
CCGCGTTGCGTCGCGGTGCATGGAGCCGGGCCACCTCGACCTGAATGGAAGCCGGCGGCACCTCG  
CTAACGGATTACCACTCCAAGAATTGGAGCCAATCAATTCTTGCGGAGAACTGTGAATGCGCAA  
ACCAACCCTTGGCAGAACATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACGCGGCGCATCTC  
GGGACGCGTTGGGTCTTGCCACGGGTGCGCATGATCGTGCTCCTGTCTGTTGAGGACCCGGCTAGG  
CTGGCGGGGTTGCCTTACTGGTTAGCAGAATGAATCACCGATACGCGAGCGAACGTGAAGCGACT  
GCTGTGCAAAACGTCTGCGACCTGAGCAACAACATGAATGGTCTTCGGTTTCCGTGTTTCGTAAA  
GTCTGGAACGCGGAAGTCAGCGCCCTGCACCATTATGTTCCGGATCTGCATCGCAGGATGCTGCT  
GGCTACCCTGTGGAACACCTACATCTGTATTAACGAAGCGCTGGCATTGACCCCTGAGTGATTTTCT  
CTGGTCCCGCCGCATCCATACCGCCAGTTGTTTACCCTCACAACGTTCCAGTAACCGGGCATGTTT  
ATCATCAGTAACCCGTATCGTGAGCATCTCTCTCGTTTCATCGGTATCATTACCCCCATGAACAG  
AAATCCCCCTTACACGGAGGCATCAGTGACCAACAGGAAAAAACCGCCCTTAACATGGCCCCG  
TTTATCAGAAGCCAGACATTAACGTTCTTGAGAACTCAACG

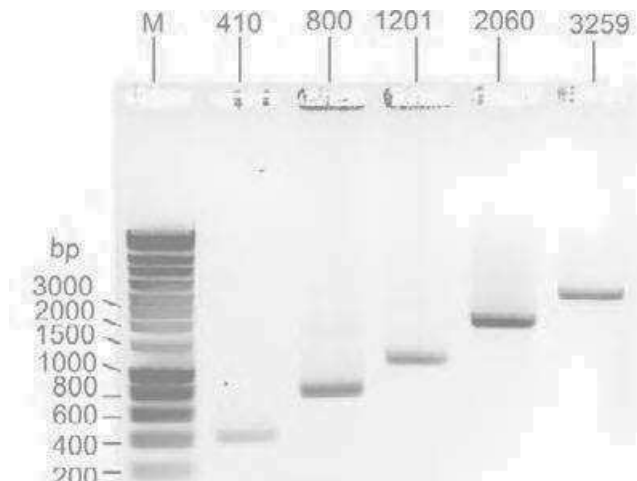
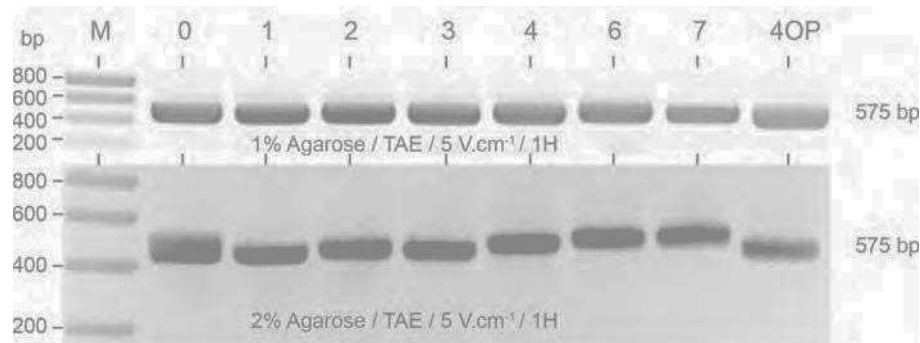


Fig. 7.34 – Gel agarose 1% coloration BET 15min de quelques un des ADNs utilisés

### B.7 Courbure - ADN 575 bp, sans et avec un nombre croissant de séquences A-tract : CA<sub>6</sub>CGG



Paramètres de la synthèse pour  $L_{ADN} = 575$  bp, sans courbure

(A) :  $L_{ADN} = 575$  bp

- Plasmide → pTOC6
- Oligo-1 → F575
- Oligo-2 → R575

Séquence 575 bp, 6An0, sans courbure

ATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGG  
GTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCAC  
ATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAG  
GCGTATCACGAGGCCCTTTCGTCTTCAAGAATTCTCATGTTTGACAGCTTATCATCGATAAGCTTCCATCT  
CGAACCGACGTTGCTGGCCGTACATGTGGCTAGCACCATGGAGGATCCCTTAACGCTATGAGAACTCGCC  
GCCCCGACTGGGCTGTGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGTCTCTTCCGGTGGGCGCGGG  
GCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGC  
GCTCTGGGTCAATTTTCGGCGAGGACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTCGCTTGC GGAT  
TCGGAATCTTGCACG

**Paramètres de la synthèse pour  $L_{ADN} = 575$  bp, avec courbure**

	6An1 <sup>P</sup>	6An2 <sup>P</sup>	6An3 <sup>P</sup>	6An4 <sup>P</sup>	6An6 <sup>P</sup>	6An7 <sup>P</sup>	6An4 <sup>O</sup>
Plasmide	pTOC7	pTOC8	pTOC9	pTOC10	pTOC12	pTOC13	pTOC11
Oligo1	F575	F575	F575	F575	F575	F575	F575
Oligo2	R575	R575	R575	R575	R575	R575	R575

**Séquence 575 bp, 6Ani<sup>P</sup> ou <sup>O</sup>, avec courbure**

Seul la constitution des l'insert de 88 bp varie en fonction de la courbure, se référer à Fig. B.7 pour le reste de la séquence.

**B.8 Courbure - ADN 575 bp, sans et avec un nombre croissant de séquences A-tract : TA<sub>3</sub>TC****Paramètres de la synthèse pour  $L_{ADN} = 575$  bp, avec et sans courbure**

	3An0	3An3 <sup>P</sup>
Plasmide	pTOC6	pTOC1
Oligo1	F575	F575
Oligo2	R575	R575

**Séquence 575 bp, 6An0, sans courbure**

Seul la constitution des l'insert de 88 bp varie en fonction de la courbure, se référer à Fig. 6.9 et à , se Fig. B.7 pour le reste de la séquence.

**B.9 Courbure - ADN 575 bp, sans et avec un nombre croissant de séquences A-tract : Rand**

L'ADN témoins, sans courbure, correspond à l'ADN 6AnO, se référer à Fig. B.7.

	Rand1 <sup>P</sup>	Rand2 <sup>P</sup>	Rand3 <sup>P</sup>
Plasmide	pTOC3	pTOC4	pTOC5
Oligo1	F575	F575	F575
Oligo2	R575	R575	R575

Seul la constitution des l'insert de 88 bp varie en fonction de la courbure, se référer à Fig. 6.10 et à , se Fig. B.7 pour le reste de la séquence.

## B.10 Dénaturation - ADN 583 bp, avec et sans coeur *polyA*



### Paramètres de la synthèse pour $L_{ADN} = 583$ bp

(A) : avec un coeur polyA de 50 bp

- Plasmide → pKH47
- Oligo-1 → O1
- Oligo-2 → O2

(B) : sans un coeur polyA

- Plasmide → pBR422
- Oligo-1 → O3
- Oligo-2 → O4

Fig. 7.35 – Gel agarose 1% coloration BET 15min.

### Séquence 583 bp *PolyA*

```
CGTTCCAGTAACCGGGCATGTTTCATCATCAGTAACCCGTATCGTGAGC
ATCCTCTCTCGTTTCATCGGTATCATTACCCCATGAACAGAAATTCCC
CCTTACACGGAGGCATCAAGTGACCAAACAGGAAAAAACGCCCTTA
ACATGGCCCGCTTTATCAGAAGCCAGACATTAACGCTTCTGGAGAAAC
TCAACGAGCTGGACGCGGATGAACAGGCAGACATCTGTGAATCGCTT
CACGACCACGCTGATGAGCTTTACCGCAGAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTGCCTCGCGCG
TTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGA
CGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGT
CAGGGCGCGTCAGCGGGTGTGGCGGGTGTGGGGGCGCAGCCATGAC
CCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACCTATGCGGCAT
CAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGC
ACAGATGCGTAAGGAGAA
```

### Séquence 583 bp

```
TACCGCCAGTTGTTTACCCTCACAACGTTCCAGTAACCGGGCATGTTT
ATCATCAGTAACCCGTATCGTGAGCATCCTCTCTCGTTTCATCGGTATC
ATTACCCCATGAACAGAAATCCCCCTTACACGGAGGCATCAGTGAC
CAAACAGGAAAAAACGCCCTTAACATGGCCCGCTTTATCAGAAGCC
AGACATTAACGCTTCTGGAGAACTCAACGAGCTGGACGCGGATGA
ACAGGCAGACATCTGTGAATCGCTTCACGACCACGCTGATGAGCTTT
ACCGCAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGA
CACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGC
CGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGG
TGTCGGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTA
TACTGGCTTAACCTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCAC
CATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCG
CATCAGGCGCTCTTCCG
```



## C SCRIPT *Mathematica* D'ANALYSE DES DONNÉES

Le premier fichier ci-dessous, correspond au *fichier bilan* automatiquement généré par la script *Mathematica* à la fin de chaque analyse de données. Ce *fichier bilan* est constitué de deux pages si nous procédons à une analyse *simple* des différents distributions de données via une seule gaussienne. Si le phénomène étudié nécessite, ou questionne, sur l'utilisation de 2 gaussiennes, alors cette analyse est réalisée en parallèle de l'analyse *simple* et le *fichier bilan* est constitué d'une troisième supplémentaire. Ainsi, les deux première pages renseigne sur l'analyse via 1 gaussienne tandis que la troisième page regroupe les étapes principales de l'analyse via 2 gaussiennes.

Le deuxième fichier correspond à l'ensemble du script *Mathematica* que nous utilisons pour analyser les données extraites lors de mesure HT-TPM.

Les fichiers d'entrées sont les *fichier.csv* créés pour chaque trajectoires à la suite de l'analyse vidéo par le logiciel *NanoMultiplex* qui contient l'ensemble des données  $X_i, Y_i, X_{anc,i}, Y_{anc,i}, F_{Sym,i}, Ae_{q,i}$  et  $Ae_{q,anc,i}$ .

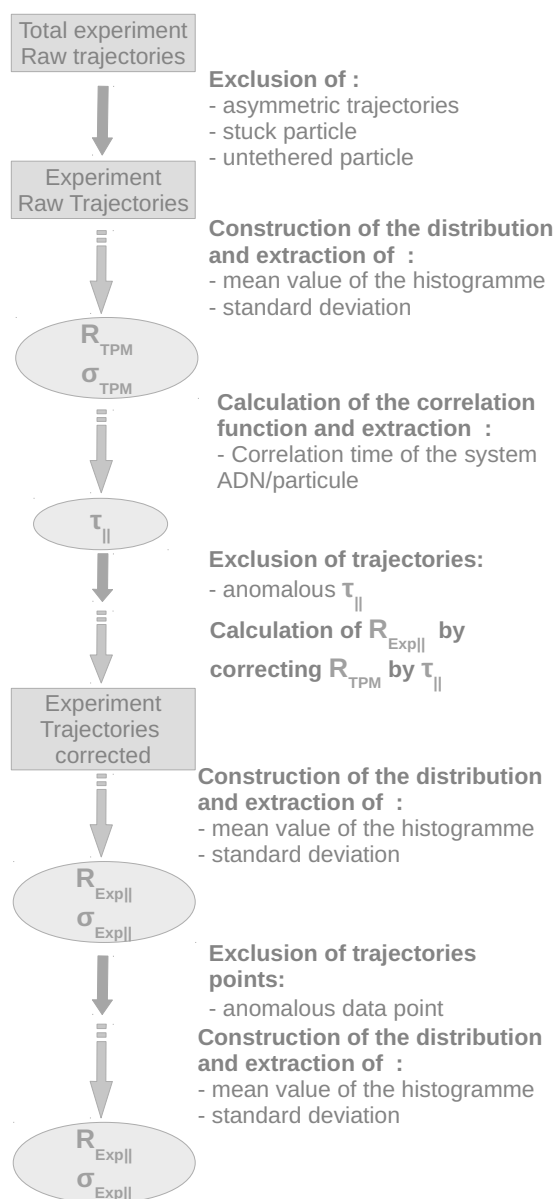
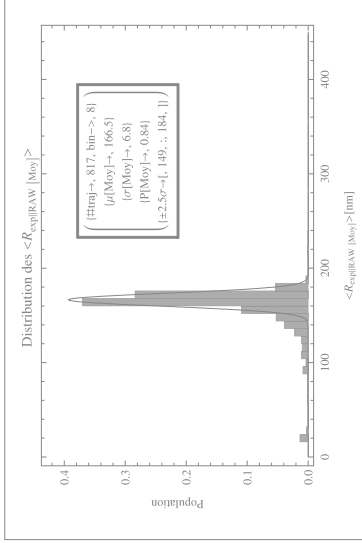


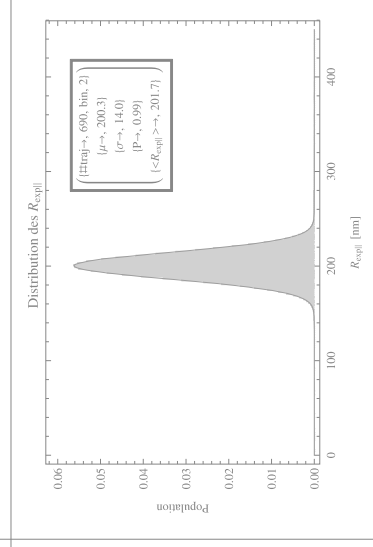
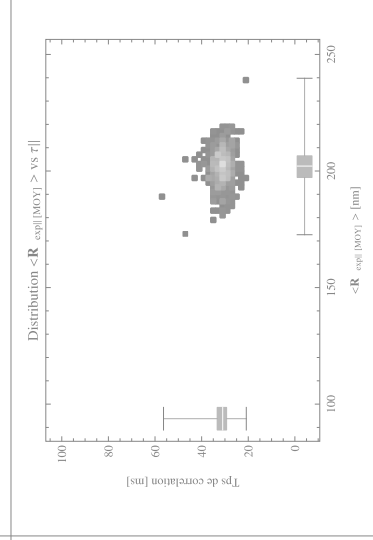
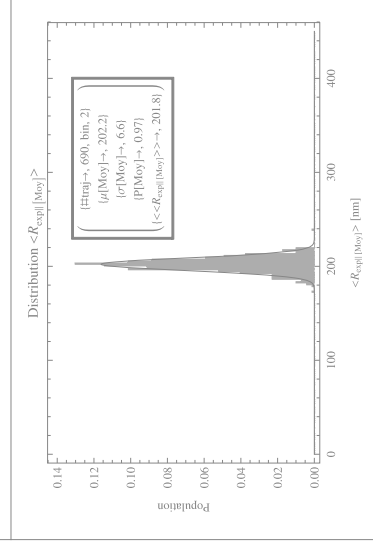
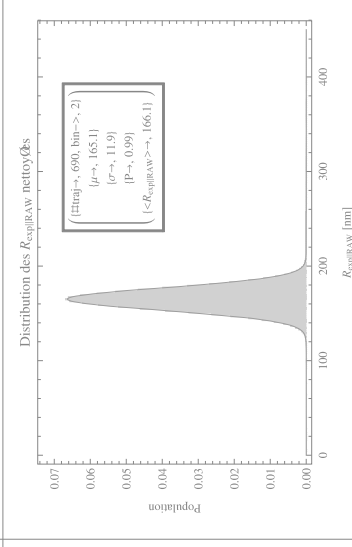
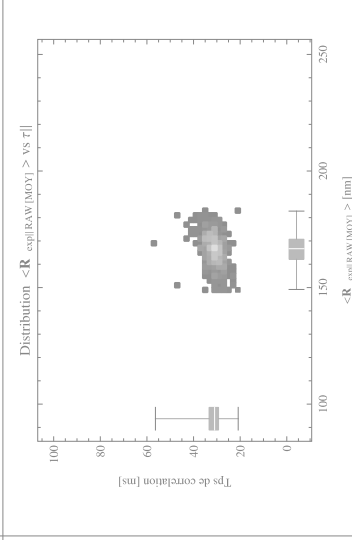
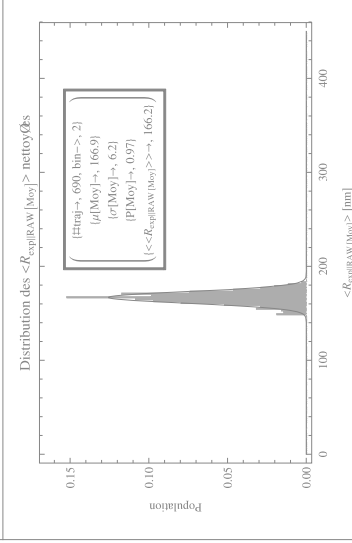
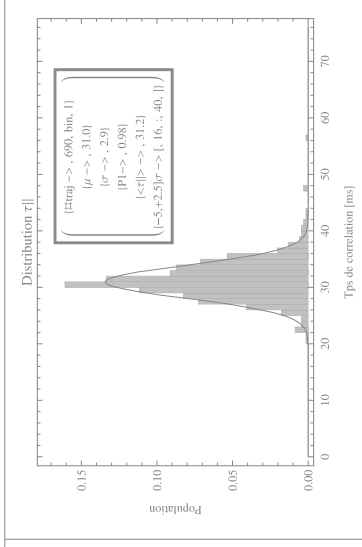
Fig. 7.36 – Logigramme simplifié du script *Mathematica*

Les fichiers de sortie sont :

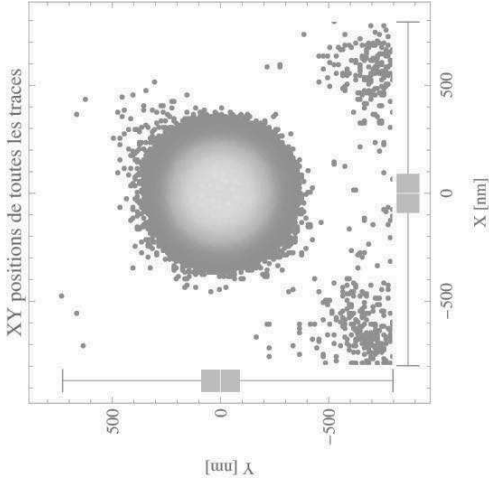
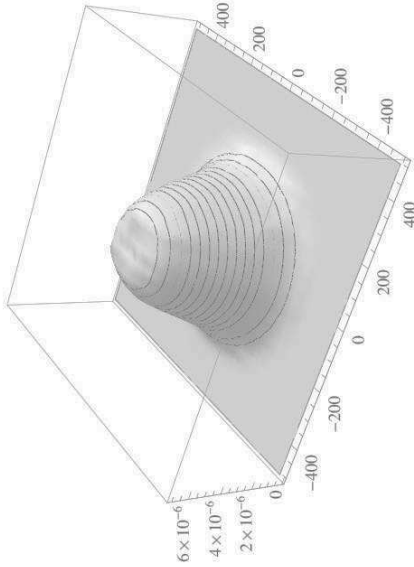
- Pour chaque étape d'exclusion un fichier bilan répertoriant les trajectoires conservées et rejetées ainsi que la valeur du paramètres d'exclusion associé.
- Suite à chaque étape de construction de distribution un fichier contenant l'ensemble des données associées à cette dernière est généré.
- Pour chaque trajectoire un fichier contenant les données du calcul des fonctions de corrélations et de la matrice de covariance associée, est généré
- Un fichier global répertoriant l'ensemble des valeurs obtenues pour chaque trajectoire et ce à chaque étape de l'analyse, est généré.
- Un *fichier bilan* répertoriant l'ensemble des représentations associées à chaque étape de l'analyse ainsi que les valeurs respectives obtenues lors de ces dernières est généré.



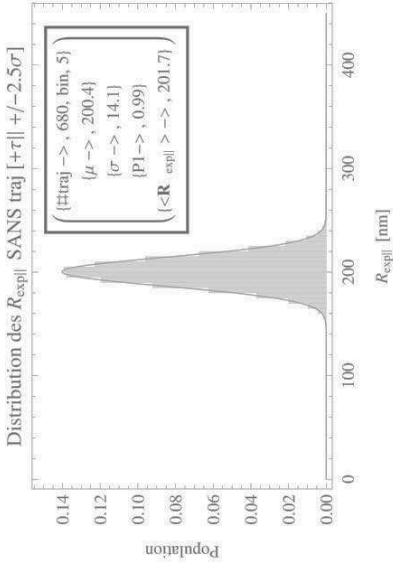
Objectif : DENATURATION  
 Manip du (Tro) sur IADN Ladh= (1201bp|bp)\_2  
 SoTampontT\_PhuL127 + NaCl 155 mM) avec pelletier (25)  
 Bilan\_A141020\_1201\_DENAT-25A\_pdf  
 AnalyseOle: 20141020  
 Page 1/2



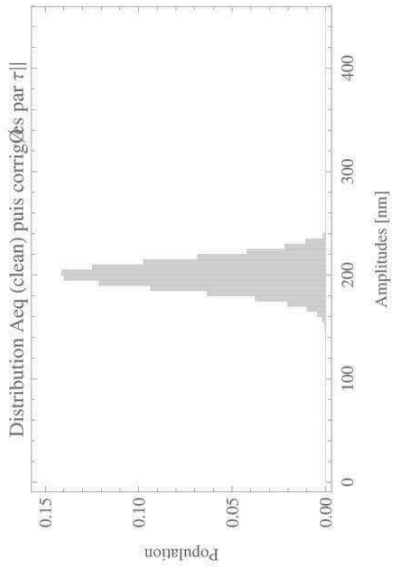
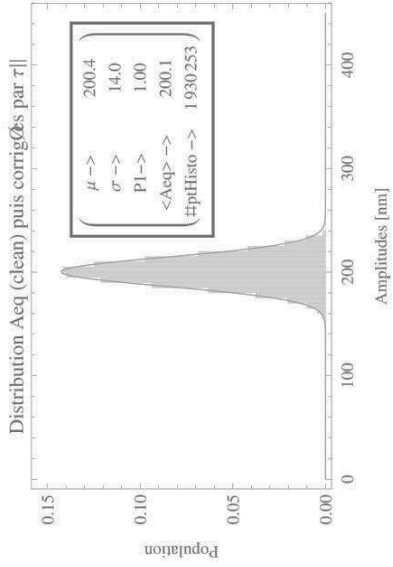
Objetif : [DENATURATION]  
Manip du [Tot] sur l'ADN Ladn= (120[bp][bp])\_2  
SolTampou[T\_PluL127 + NaCl 155 mM] avec pelier (25)  
Bilan\_A141020\_1201\_DENAT-25A\_.pdf  
Analyse (20141020)  
Page 2/2



Nombre totale de traces analysées : [680]  
Valeur Fit-G [Rexp][RAW] [nm] : [165.1]  
Valeur Fit-G [Rexp] [nm] : [200.4]  
Valeur Mean [Rexp] [nm] : [201.7]  
Valeur Fit-G [rll] [nm] : [31.0]  
Valeur Mean [rll] [nm] : [31.2]  
Taille des bins [nm] : [5]



SUPPRESSION des points tq :  
Fit-G [Rexp] - 5 sigma < Rexp < Fit-G [Rexp] + 2.5 sigma :  
=> {130} [nm] < Rexp < {236} [nm]  
=> #Point Initial : {1962211}  
=> #Point Supprimé : {31957}  
=> #Point Final : {1930253}  
Valeur Fit-G [Rexp] [nm] : [200.38]  
Valeur Mean [Rexp] [nm] : [200.11]  
Taille des bins [nm] : [5]



Première analyse : [680]  
Valeur Fit-G [Rexp][RAW] [nm] : [165.1]  
Valeur Fit-G [Rexp] [nm] : [200.4]  
=> RexpADN [BASIC] [nm] : [194.3]  
=> LpWLC [RexpADN [BASIC]] [nm] : [{}]  
=> RexpADN [Segall] [nm] : [196.2]  
=> LpWLC [RexpADN [Segall]] [nm] : [{}]

## Programme d'Analyse des données TPM

-- PRÉAMBULE --

Que fait ce programme ?

```
(* Programme d'analyse des données TPM *)

(* Input : fichier.csv générés par le logiciel Nanomultiplex (TPM) *)
(* Analyse ;;
    - Selection des trajectoires valides;
    - Extraction de  $R_{TPM}$ ;
    - Correction de l'erreur de pointée;
    - Extraction du temps de corrélation  $\tau_{||}$ ;
    - Correction du l'effet de moyennage vidéo, lie à  $\tau_{||}$ ;
    - Extraction de  $R_{Exp||}$ ;)

(* Output ;;
    - fichier echo_Fsym.tex, echo_traj.tex,
    echo_ampl.tex, fichiers bilan de chaque étape d'exclusion
    de trajectoires;

    - fichier de calcul des fonctions et matrices de
    corrélations : MatCorr.dat =
    {Xcor(.csv), X(i)*X(i+j) [...] Ycor(.csv), Y(i)*Y(i+j)}
    avec  $j \in [0:taumax]$  Et Tcorr.dat = {Tex*i, Cm(t)/Cm(0)};
    - fichier histro_etape(i).dat,
    fichiers contenant l'ensemble des données de chaque
    distribution de probabilité;
    - fichier InfoFile.dat, répertoriant l'ensemble
    des valeurs obtenues pour chaque
    trajectoire et à chaque étape de l'analyse;
    - fichier "bilan", répertoriant l'ensemble
    des représentations associées à
    chaque étape de l'analyse ainsi que les valeurs respectives obtenues lors
    de ces dernières;*)
```

Sélection du repertoire de travail

```
(* Nettoyage des parametres globaux *)
ClearAll["Global'*"]
ClearAll[H1, SubText1, HDtau, H3, davantcorr, H2, H5, daprescorr, H4,
SubText2, pdensity2, pdensity1, SubText3, HExp, SubText4, SubText5,
HNew, HNewbis]
(* ===== Selection de l'arborescence du fichier ===== *)
SetDirectory["/home/brunet/Bureau/Denat_1201_25/"];
directory = Directory[];
```

-- TRAITEMENT DES DONNÉES **BRUTES** --

Copie des scripts shell

```
(*Commande shell pour copier les scripts dans le repertoire courant*)
cmd1 = "cp /home/brunet/Documents/shell/shell/aa_fsym_v-universelle.sh .";
cmd2 = "cp /home/brunet/Documents/shell/shell/aa_traj_v-universelle.sh .";
cmd3 = "cp /home/brunet/Documents/shell/shell/aa_ampl_v-universelle.sh .";
Run[cmd1, cmd2, cmd3];
```

## Communication des variable aux scripts shell

```

(* ===== Indice du programme ===== *)
(* i dans InfoFile : nbr de complexes ADN/particule;
   j dans boucle d'écriture : nbr de complexes ADN/particule;
   n taille des données Xcor et Ycor*)

(* ===== Paramètres globaux ===== *)
a0 = 0.34; (*Unité de conversion = 1bp = 0.34 nm*)
ErrP = 15.0; (*en nm. Erreur de pointé =
 $\sqrt{\langle R^2_{\text{TPM}} \rangle}$  [D'une particule fixe] [nm] *)

(* ===== Paramètres globaux spécifique à la manip ===== *)
DateManip = ToString["Tot_"];
ADNManip = ToString["1201-"];
ParametreManip = ToString["T-25_"];
ParticulariteManip =
  ToString[StringJoin[{DateManip, ADNManip, ParametreManip}]]];

L = 1201 * 0.34; (*Longueur de l'ADN = #bp*a0=nm*)
Rpar = 150.0; (*Rayon particule en nm*)
TauCaracADN = 37; (*ms*)
Tex = 0.04; (*s → 40ms*)
Tac = 0.04; (*s → 40ms*)
Tav = 2.; (*s*)
(* =====
   Paramètres globaux spécifique a la version du logiciel ===== *)
nCSV = 52; (*Interation/indice du fichier .CVS à partir duquel les
valeurs sont ecrites*)
nPtFichierCSV = 2 * 60 * (1 / Tac);
(*Taille des fichier de données pour allouée la memoire*)
(* Renseignement des index des lignes des fihcier.cvs générés par
le logiciel NanoMultiplex *)
(*(* OLD version de logiciel *)
NligneX=1;
NligneXanc=3;
NligneXcor=0;
NligneY=2;
NligneYanc=4;
NligneYcor=0;
NligneFSym=10;
NligneAeq=11;*)

(* OLD version de Germain *)
NligneX = 1;
NligneXanc = 3;
NligneXcor = 0;
NligneY = 2;
NligneYanc = 4;
NligneYcor = 0;
NligneFSym = 5;
NligneAeq = 6;

```

## Permiers FILTRES de Sélection

1er FILTRE de Sélection : -> Filtre sur **F<sub>Sym</sub>** (FACTEUR DE SYMÉTRIE)

Lancement 1 er script shell :

```
(* Commande de lancement des script shell
avec Modification de la syntaxe pour etre lu par Mathematica
Ssi <FSym>[Moy] (== moyenne sur l'ensemble des FSym de la trajectoire)>
1.35 → traces rejetées *)
cmdfsym = "sh aa_fsym_v-universelle.sh " <> ToString[nCSV] <> " " <>
ToString[NligneFSym] <> " " <> ToString[NligneAeq] <> " " >echo_fsym.txt";
Run[cmdfsym];
```

```
DataNbrFsym = Import["echo_fsym.txt", "Table"];
Dimensions[DataNbrFsym];
NdEndFS = DataNbrFsym[[Length[DataNbrFsym]]];
NbrFsym = NdEndFS[[3]];
Print["# trajctaires validés via : <FSym> ≤ 1.35 : ", NbrFsym]
```

# trajctaires validés via : <F<sub>Sym</sub>> ≤ 1.35 : 912

2eme FILTRE de Sélection : -> Filtre sur **R<sub>exp||Raw</sub>**(i) : trajectoires qui se collent ou se décrochent

Lancement 2 éme script shell :

```
(* Ssi ;;
- Rexp||Raw(i), pour i qcq > 1000 nm → traces rejetées;
- Rexp||Raw(i), pour i qcq < 1 nm → traces rejetées; *)
cmdtraj = "sh aa_traj_v-universelle.sh " <> ToString[nCSV] <> " " <>
ToString[NligneAeq] <> " " >echo_traj.txt";
Run[cmdtraj];
(*Extration du nombre de trajectoire valides à la suite de cette étape*)
DataNbrTraj = Import["echo_traj.txt", "Table"];
Dimensions[DataNbrTraj];
NdEnd2 = DataNbrTraj[[Length[DataNbrTraj] - 1]];
NbrTraj = NdEnd2[[3]];
Print["# trajctaires validées via : 1 nm < R(i) < 1000 nm : ",
NbrTraj]
```

# trajctaires validées via : 1 nm < R(i) < 1000 nm : 881

3eme FILTRE de Sélection : -> Filtre sur <R<sub>exp||Raw</sub>> (AMPLITUDE MOYENNEE SUR TOUTE LA DURÉE DE LA TRAJECTOIRE)

```
(* Ssi : - <Rexp||Raw Moy> ∉ <Rexp||Raw Moy> ±2.5σ → traces rejetées; *)
```

Représentation/histo pour définir les bornes du filtre

```
(* ===== Histo ===== *)
NomHistoMoy = "histoAeqMoy.dat";
NomArboHistoMoy = ToFileName[directory, NomHistoMoy];
(*Nom du fichier output*)
DataMoy = Import[NomArboHistoMoy, "Table"];
(* Variable de construction des histogrammes*)
Amax = 500;
```

```

Amin = 0;
AeqMinTraj = Min[DataMoy];
AeqMaxTraj = Max[DataMoy];
Print["Taille du bin = ",
      sb = Ceiling[((AeqMaxTraj - AeqMinTraj) + 10) / Sqrt[NbrTraj]], "[nm]"];
If[sb > 15.0,
  Print[Style["== ATTENTION ==", 15, Red]];
  Print["MAIS bins mit à 15 [nm]"];
  sb = 15.0;
]
(*le "+10" qui intervient dans la définition de la taille du bin
est décidé arbitrairement*)
(* RMQ : Ici, nous nous interessons seulement aux amplitudes BRUTES
MOYENNES. C  d les RTPM sont moy  nn  s sur tout la trajectoire afin
d'obtenir UNE SEULE valeur par trace de complexe ADN/particule. Donc
pour UNE trajectoire nous avons UNE valeur moyennes. Ce premier
histogramme est constitu   de N valeurs moyennes de  $\sqrt{\langle R^2_{TPM} \rangle}$ 
correspondant    nos N traces. Ainsi, la taille des bin est grande,
car, ici, nous consid  rons peu de variable independantes. Normalement sizebin=
 $\frac{AeqMax-AeqMin}{\sqrt{N}}$  *)

(* Donn  es des amplitudes Moyenn  es sur le Tac    AeqMoy *)
(*Contruction de l'histogramme*)
lgrhistoMoy = Length[DataMoy];
DataMoy = Table[DataMoy[[i]][[1]], {i, 1, lgrhistoMoy}];
Histogram[DataMoy, {Amin, Amax, sb}, "Probability"];
(*Contruction de l'ajustement*)
{bins, counts} = HistogramList[DataMoy, {Amin, Amax, sb}, "Probability"];
centers = MovingAverage[bins, 2];

modelMoy = sb * 
$$\left( (P1) * \frac{e^{-\frac{(x-\mu)^2}{2 \sigma^2}}}{\sqrt{2 \pi} \sigma} \right);$$


parsMoy = FindFit[{centers, counts}^T, modelMoy, {{  , 130}, {  , 10}, {P1, 0.8}},
  x];
(*Variable extraites pour la suite*)
MeanMoy =    /. parsMoy[[1]];
SDMoy =    /. parsMoy[[2]];
Print[
  "Bornes Ssi application du 3eme filtre :  $\mu[R_{exp}|RAW [Moy]] - 5\sigma <$ 
 $<R_{exp}|Raw Moy > < \mu[R_{exp}|RAW [Moy]] + 2.5\sigma \rightarrow$ [" ,
  AmpMin = Floor[MeanMoy - 2.5 * SDMoy], ":", AmpMax = Ceiling[MeanMoy + 2.5 * SDMoy],
  "]" nm"];
Print[" $\mu[R_{exp}|RAW [Moy]]$  (FitG) = ", SetAccuracy[MeanMoy, 2], " nm"];
Print[" $<R_{exp}|RAW [Moy]>$  (Mean) = ", SetAccuracy[Mean[DataMoy], 2], " nm"];
(*Trac  *)
H1 = Show[Histogram[DataMoy, {Amin, Amax, sb}, "Probability",
  PlotLabel    "Distribution des  $<R_{exp}|RAW [Moy]>$  ", Frame    True,
  FrameLabel    {" $<R_{exp}|RAW [Moy]>$  [nm]", "Population"},
  ChartLabels   
  {Placed[
    Framed[MatrixForm[{{"#traj->", lgrhistoMoy, "bin->", sb},

```

```

{"μ[Moy]→", SetAccuracy[μ /. parsMoy[[1]], 2]},
{"σ[Moy]→", SetAccuracy[σ /. parsMoy[[2]], 2]},
{"P[Moy]→", SetAccuracy[P1 /. parsMoy[[3]], 3]},
{"±2.5σ→[" , AmpMin, ":", AmpMax, "]" }]]], {0.7, 0.8}]],
ChartStyle → Gray], Plot[modelMoy /. parsMoy // Evaluate, {x, Amin, Amax},
PlotRange → All, PlotStyle → Black]]

```

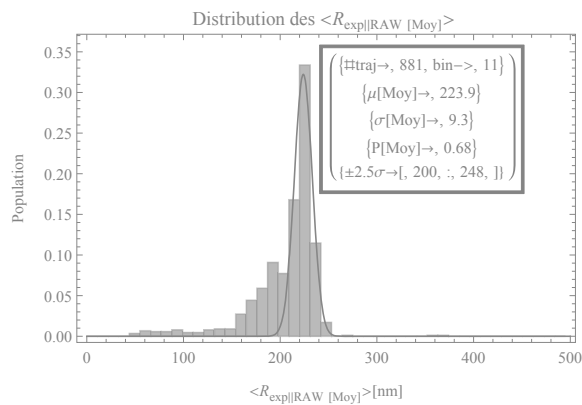
Taille du bin = 11[nm]

Bornes Ssi application du 3eme filtre :

$$\mu[R_{\text{exp}} | \text{RAW [Moy]}] - 5\sigma < \langle R_{\text{exp}} | \text{Raw Moy} \rangle < \mu[R_{\text{exp}} | \text{RAW [Moy]}] + 2.5\sigma \rightarrow [200:248] \text{ nm}$$

$$\mu[R_{\text{exp}} | \text{RAW [Moy]}] (\text{FitG}) = 223.9 \text{ nm}$$

$$\langle R_{\text{exp}} | \text{RAW [Moy]} \rangle (\text{Mean}) = 206.4 \text{ nm}$$





## SSi 2 populations -- Représentation de l'ensemble des points d'amplitude

```

(* ===== Histo ===== *)

Histogram[DataMoy, {Amin, Amax, sb}, "Probability"];
{bins, counts} = HistogramList[DataMoy, {Amin, Amax, sb}, "Probability"];
centers = MovingAverage[bins, 2];

modelMoy2G = sb * 
$$\left( (P1) * \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} + (P2) * \frac{e^{-\frac{(x-\mu2)^2}{2\sigma2^2}}}{\sqrt{2\pi}\sigma2} \right);$$


parsMoy2G = FindFit[{centers, counts}^T, modelMoy2G,
  {{μ, 180}, {σ, 10}, {P1, 0.5}, {μ2, 205}, {σ2, 10}, {P2, 0.5}}, x];

(*Variable extraites pour la suite*)
MeanMoy2G1 = μ /. parsMoy2G[[1]];
MeanMoy2G2 = μ2 /. parsMoy2G[[4]];
SDMoy2G1 = σ /. parsMoy2G[[2]];
SDMoy2G2 = σ2 /. parsMoy2G[[5]];

Print["μ[Rexp||RAW [Moy]] (FitG) = ", SetAccuracy[MeanMoy2G1, 2], " nm",
  SetAccuracy[MeanMoy2G2, 2], " nm"];
Print["<Rexp||RAW [Moy]> (Mean) = ", SetAccuracy[Mean[DataMoy], 2], " nm"];

(*Tracé*)
H12G = Show[Histogram[DataMoy, {Amin, Amax, sb}, "Probability",
  PlotLabel → "Distribution des <Rexp||RAW [Moy]> ", Frame → True,
  FrameLabel → {"<Rexp||RAW [Moy]>[nm]", "Population"},
  ChartLabels →
    {Placed[
      Framed[MatrixForm[{{"#traj→", lgrhistoMoy, "bin→", sb},
        {"μ[Moy]→", SetAccuracy[μ /. parsMoy2G[[1]], 2],
          SetAccuracy[μ2 /. parsMoy2G[[4]], 2}],
        {"σ[Moy]→", SetAccuracy[σ /. parsMoy2G[[2]], 2],
          SetAccuracy[σ2 /. parsMoy2G[[5]], 2}],
        {"P[Moy]→", SetAccuracy[P1 /. parsMoy2G[[3]], 3],
          SetAccuracy[P2 /. parsMoy2G[[6]], 3]]}], {0.8, 0.8}],
  ChartStyle → Gray], Plot[modelMoy2G /. parsMoy2G // Evaluate,
  {x, Amin, Amax}, PlotRange → All, PlotStyle → Black]]

```

## SSI 1 population -- Représentation de l'ensemble des points d'amplitude

```
(* ===== Histo ===== *)
NomHisto = "histoTraj.dat"; (*Nom du fichier output*)
NomArboHisto = ToFileName[directory, NomHisto];
Data = Import[NomArboHisto, "Table"];
(*Données ensemble des valeurs d'amplitudes du mouv → Aeq *)
lgrhisto = Length[Data];
Data = Table[Data[[i]][[1]], {i, 1, lgrhisto}];
Histogram[Data, {Amin, Amax, sb}, "Probability"];
{bins, counts} = HistogramList[Data, {Amin, Amax, sb}, "Probability"];
centers = MovingAverage[bins, 2];

model = sb * 
$$\left( P1 * \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \right);$$

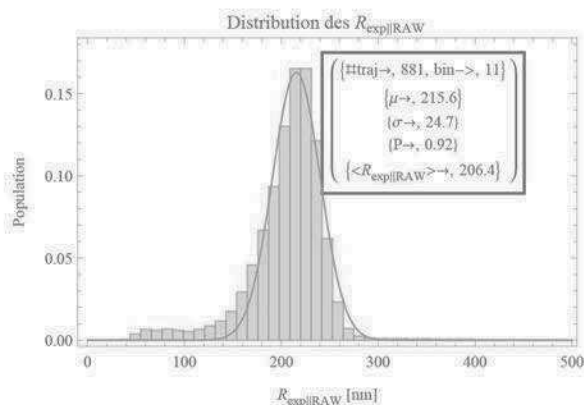

pars = FindFit[{centers, counts}^T, model, {{μ, 130}, {σ, 8}, {P1, 0.5}}, x];
(* Ré definition des variable de construction de l'histo*)
(*Amax=Ceiling[(μ/.pars[[1]])+(0.12*(σ/.pars[[2]]))*(μ/.pars[[1]])];
sbhistol=(Ceiling[(μ/.pars[[1]])+2.5((σ/.pars[[2]])/(nbrparl^(1/2)))]);*)

Print["μ[Rexp||RAW] (FitG) = ", MeanMoy = μ /. pars[[1]], " nm"];
Print["<Rexp||RAW> (Mean) = ", SetAccuracy[Mean[Data], 2], " nm"];

H1bis = Show[Histogram[Data, {Amin, Amax, sb}, "Probability",
  PlotLabel → "Distribution des Rexp||RAW", Frame → True,
  FrameLabel → {"Rexp||RAW [nm]", "Population"},
  ChartLabels →
    {Placed[
      Framed[MatrixForm[{{{"#traj→", lgrhistoMoy, "bin→", sb},
        {"μ→", SetAccuracy[μ /. pars[[1]], 2]},
        {"σ→", SetAccuracy[σ /. pars[[2]], 2]},
        {"P→", SetAccuracy[P1 /. pars[[3]], 3]},
        {"<Rexp||RAW>→", SetAccuracy[Mean[DataMoy], 2]}]}], {0.7, 0.8}]]],
  Plot[model /. pars // Evaluate, {x, Amin, Amax}, PlotRange → All]]
```

$\mu[R_{\text{exp}}||\text{RAW}]$  (FitG) = 215.596 nm

$\langle R_{\text{exp}}||\text{RAW} \rangle$  (Mean) = 206.4 nm



## SSi 2 populations -- Représentation de l'ensemble des points d'amplitude

Lancement 3<sup>ème</sup> script shell :

```
cmdampl = "sh aa_ampl_v-universelle.sh " <> ToString[nCSV] <> " " <>
  ToString[AmpMax] <> " " <> ToString[AmpMin] <> " " <> ToString[NligneAeq]
  " >echo_ampl.txt";
Run[cmdampl];
(*Extraction du nombre de trajectoire valides à la suite de cette étape*)
DataNbrAmpl = Import["echo_ampl.txt", "Table"];
Dimensions[DataNbrAmpl];
NdEnd3 = DataNbrAmpl[[Length[DataNbrAmpl]]];
NbrAmpl = NdEnd3[[3]];
Print[
  "# trajectoires validées via :  $\mu[R_{exp}|RAW_{Moy}] - 5\sigma < <R_{exp}|Raw Moy >$ 
  <  $\mu[R_{exp}|RAW_{Moy}] + 2.5\sigma$  : ", NbrAmpl]
```

# trajectoires validées via :

$\mu[R_{exp}|RAW_{Moy}] - 5\sigma < <R_{exp}|Raw Moy > < \mu[R_{exp}|RAW_{Moy}] + 2.5\sigma$  : 612

## Vérification par représentation/histo + Écriture du fichier histoGlobal.dat

```
(* =====!!! Histo CLEAN !!!===== *)
NomHisto2 = ToString[StringForm["histoSup``.dat", (AmpMin)]];
(*Nom du fichier*)
NomArboHisto2 = ToFileName[directory, NomHisto2];(*Chemin*)
Data2 = Import[NomArboHisto2, "Table"];
(*Importation du fichier selon le chemin *)
lgrhisto2 = Length[Data2];

NomHistoMoy2 = ToString[StringForm["histoAeqMoySup``.dat", (AmpMin)]];
NomArboHistoMoy2 = ToFileName[directory, NomHistoMoy2];
DataMoy2 = Import[NomArboHistoMoy2, "Table"];
lgrhistoMoy2 = Length[DataMoy2];
nbrpar1 = lgrhistoMoy2;

(* Variable de construction des histogrammes*)
Amax = 400;
Amin = 0;
nbrpar1 = 100;
Print["Taille du bin = ",
  sb2 = Ceiling[((AmpMax - AmpMin) + 10) / Sqrt[(NbrAmpl)]]];
If[sb2 > 15.0,
  Print[Style["== ATTENTION ==", 15, Red]];
  Print["MAIS bins mit à 15 [nm]"];
  sb2 = 15.0;
  ] (*sb2=Ceiling[150/(nbrpar1^(1/2))]; Ici,
  taille moyenne  $\rightarrow \frac{AeqMax-AeqMin}{\sqrt{N}} = \frac{100}{\sqrt{200}} \sim 7 \text{ nm}$ *)

(* CLEAN Données des amplitudes Moyennées sur le Tac  $\rightarrow$  AeqMoy *)
DataMoy2 = Table[DataMoy2[[i]][[1]], {i, 1, lgrhistoMoy2}];
Histogram[DataMoy2, {Amin, Amax, sb2}, "Probability"];
{bins, counts} = HistogramList[DataMoy2, {Amin, Amax, sb2}, "Probability"];
centers = MovingAverage[bins, 2];
```

```

modelMoy2 = sb2 * 
$$\left( (P1) * \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \right);$$


parsMoy2 = FindFit[{centers, counts}^T, modelMoy2,
  {{μ, 200}, {σ, 10}, {P1, 0.6}}, x];
H3 = Show[Histogram[DataMoy2, {Amin, Amax, sb2}, "Probability",
  PlotLabel → "Distribution des <Rexp||RAW [Moy]> nettoyyées", Frame → True,
  FrameLabel → {"<Rexp||RAW [Moy]> [nm]", "Population"},
  ChartLabels →
    {Placed[
      Framed[MatrixForm[{{"#traj→", lgrhistoMoy2, "bin→", sb2},
        {"μ[Moy]→", SetAccuracy[μ /. parsMoy2[[1]], 2}},
        {"σ[Moy]→", SetAccuracy[σ /. parsMoy2[[2]], 2}},
        {"P[Moy]→", SetAccuracy[P1 /. parsMoy2[[3]], 3}},
        {"<Rexp||RAW [Moy]>→", SetAccuracy[Mean[DataMoy2], 2]}}]],
      {0.25, 0.8}]], ChartStyle → Gray],
  Plot[modelMoy2 /. parsMoy2 // Evaluate, {x, Amin, Amax}, PlotRange → All,
  PlotStyle → Black]

(* CLEAN Données ensemble des valeurs d'amplitudes du mouv → Aeq *)
Data2 = Table[Data2[[i]][[1]], {i, 1, lgrhisto2}];
Histogram[Data2, {Amin, Amax, sb2}, "Probability"];
{bins, counts} = HistogramList[Data2, {Amin, Amax, sb2}, "Probability"];
centers = MovingAverage[bins, 2];

model2 = sb2 * 
$$\left( (P1) * \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \right);$$


pars2 = FindFit[{centers, counts}^T, model2, {{μ, 200}, {σ, 10}, {P1, 0.6}}, x];
H2 = Show[Histogram[Data2, {Amin, Amax, sb2}, "Probability",
  PlotLabel → "Distribution des Rexp||RAW nettoyyées", Frame → True,
  FrameLabel → {"Rexp||RAW [nm]", "Population"},
  ChartLabels →
    {Placed[
      Framed[MatrixForm[{{"#traj→", lgrhistoMoy2, "bin→", sb2},
        {"μ→", SetAccuracy[μ /. pars2[[1]], 2}},
        {"σ→", SetAccuracy[σ /. pars2[[2]], 2}},
        {"P→", SetAccuracy[P1 /. pars2[[3]], 3}},
        {"<Rexp||RAW>→", SetAccuracy[Mean[Data2], 2]}}]],
      {0.8, 0.8}]],
  Plot[model2 /. pars2 // Evaluate, {x, Amin, Amax}, PlotRange → All]

(* ===== Écriture du fichier GLOBAL des Histogramme ===== *)

(Donnee = Table[0.0, {10}, {4}];
(*Écriture des entêtes*)
Donnee[[1, 1]] = Directory[];
Donnee[[1, 2]] = ToString[Date[]];
Donnee[[1, 3]] = lgrhistoMoy2;
Donnee[[1, 4]] = ParticulariteManip;
```

```

Donnee[[2, 1]] = ToString["Aeq"];
Donnee[[2, 2]] = ( $\mu$  /. pars[[1]]);
Donnee[[2, 3]] = ToString["AeqMoy"];
Donnee[[2, 4]] = ( $\mu$  /. parsMoy[[1]]);

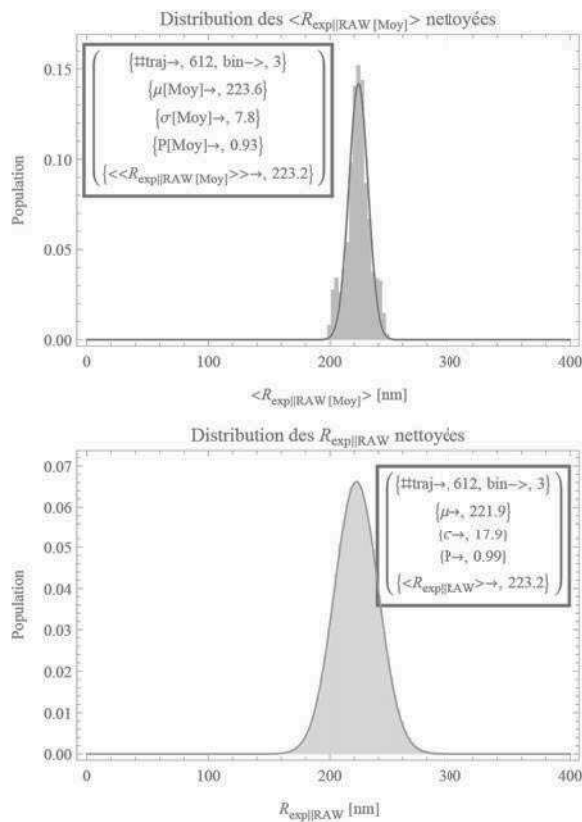
Donnee[[3, 1]] = ToString["AeqClean"];
Donnee[[3, 2]] = ( $\mu$  /. pars2[[1]]);
Donnee[[3, 3]] = ToString["AeqMoyClean"];
Donnee[[3, 4]] = ( $\mu$  /. parsMoy2[[1]]);

donnee = SetAccuracy[Donnee, 6];

(*=====*)
Export["DataGlobal.dat", donnee, "Table"];
(*=====*)
);

```

Taille du bin = 3



SSI 2 populations -- Représentation de l'ensemble des points d'amplitude

-- TRAITEMENT DES DONNÉES **CORRIGÉES** de  $\tau_{||}$  --

---

Calcul de  $\tau_{||}$  & correction des données TPM

Calcule fonction de corrélations

Initialisation des parametres pour le calcul de  $\tau_{||}$  :

```

(* ===== Écriture du fichier GLOBAL des résultats ===== *)
Fichier = FileNames["*.csv"];
DimF = Dimensions[Fichier];
nPtFichierCSV = 2 * 60 * 25;

(*Pour correction des AeqMOYClean*)
InfoFile = Table[0.0, {(DimF[[1]] + 2)}, {13}];
histoAeqMoyCleanTcor = Table[0.0, {(DimF[[1]] + 2)}, {1}];
histoAeqMoyClean = Table[0.0, {(DimF[[1]] + 2)}, {1}];
TpsCorrelation = Table[0.0, {(DimF[[1]] + 2)}, {1}];
TpsTaumax = Table[0.0, {(DimF[[1]] + 2)}, {1}];

(*Pour correction des AeqClean*)
MeanAeqCorr = Table[0.0, {(DimF[[1]] + 2)}, {2}];
InfoFile[[1, 1]] = Directory[]; (*Répertoire*)
InfoFile[[1, 2]] = ToString[DimF]; (*Nombre de fichiers*)
InfoFile[[1, 3]] = ToString[Date[]]; (*Date de traitement*)
InfoFile[[1, 4]] = ParticulariteManip; (*Rmq sur manip*)
InfoFile[[2, 1]] = ToString["(*Nom_Du_Fichier_Traite)"];
InfoFile[[2, 2]] = ToString["N°Complex"];
InfoFile[[2, 3]] = ToString[" $\tau$ "];
InfoFile[[2, 4]] = ToString[" $\Delta$ taumax"];
InfoFile[[2, 5]] = ToString["AeqMoyClean"];
InfoFile[[2, 6]] = ToString["AeqMoyC $\tau$ "];
InfoFile[[2, 7]] = ToString["AeqC $\tau$ |Moy"];
InfoFile[[2, 8]] = ToString["N°Complex"];
InfoFile[[2, 9]] = ToString["Indice_Analyse"];
InfoFile[[2, 10]] = ToString[" $\tau$ |->Oui/Non"];

(*Extraction et Écriture des numéro des fichiers de Particule/DNA*)
(*Listés dans l'ordre d'apparition/lecture machine*)
Export["InfoFile.dat", Fichier, "Table"];
NumFile = ReadList["InfoFile.dat", Word,
  WordSeparators -> {"a_a_a_ResultsParticle", ".csv"}];

(*Écriture des noms de fichiers de Particule/DNA*)
(* et ecriture des numéros des fichiers*)
For[i = 3, i < (DimF[[1]] + 3), i++,
  InfoFile[[i, 1]] = Fichier[[i - 2]];
  InfoFile[[i, 2]] = NumFile[[i - 2]];
];
(*InfoFile//TableForm*)
Export["InfoFile.dat", InfoFile, "Table"];

(*Pour ecriture des Rexp*)
EcritureRexp = Table["Rexp", {(nPtFichierCSV)}];
cptFileUsed = 0;

```

Calcul de  $\tau_{||}$  & correction de  $\mathbf{R}_{exp||Raw}$ :

```

For[j = 3, j < (DimF[[1]] + 3), j++,
  cptFileUsed = cptFileUsed + 1;
  NumParticule = InfoFile[[j, 2]];
  NomFichier = InfoFile[[j, 1]];

```

```

NomArbo = ToFileName[directory, NomFichier];

(* ===== Extraction des fichiers .cvs ===== *)
table = Import[NomArbo, "Table"];
Dim = Dimensions[table];
Dimmat = Dimensions[mat];
lgr = (Dim[[1]] - nCSV + 1);
(*Longueurs des valeurs correctes pour Xcor et Ycor*)
taumax = (11 + 1);
n = (2 + 2 * taumax);

(* ===== Calcul des fonctions de corrélations ===== *)
mat = Table[0.0, {lgr}, {n}];
(*Écriture des entêtes*)
mat[[1, 1]] = ToString["#Xcor"] (*Part[table[[1]],5]*);
mat[[1, (taumax + 2)]] = ToString["Ycor"]; (*Part[table[[1]],6]*);

For[ent = 2, ent < taumax + 2, ent++,
  mat[[1, ent]] = ToString[StringForm["C`~tX", (ent - 2)]];
  mat[[1, ((taumax + 1) + ent)]] = ToString[StringForm["C`~tY", (ent - 2)]];
];

(*Écriture des X(i)*X(i+j) et Y(i)*Y(i+j) avec j∈[0:taumax]*)
(*      pour le caclul des fonctions de corrélation      *)
For[i = 2, i ≤ (lgr), i++,

  (* Calcule de Xcor et Y cor*)
  mat[[i, 1]] = ((Part[table[[i + (26 - 2)]]], NligneX)) -
    (Part[table[[i + (26 - 2)]]], NligneXanc));
  mat[[i, (taumax + 2)]] =
    ((Part[table[[i + (26 - 2)]]], NligneY)) -
    (Part[table[[i + (26 - 2)]]], NligneYanc));

  For[ent = 2, ent < taumax + 2, ent++,
    mat[[i, ent]] =
      ((Part[table[[i + (26 - 2)]]], NligneX)) -
      (Part[table[[i + (26 - 2)]]], NligneXanc)) *
      ((Part[table[[i + (26 - 2) + (ent - 2)]]], NligneX)) -
      (Part[table[[i + (26 - 2) + (ent - 2)]]], NligneXanc));
    mat[[i, ((taumax + 1) + ent)]] =
      ((Part[table[[i + (26 - 2)]]], NligneY)) -
      (Part[table[[i + (26 - 2)]]], NligneYanc)) *
      (Part[table[[i + (26 - 2) + (ent - 2)]]], NligneY) -
      (Part[table[[i + (26 - 2) + (ent - 2)]]], NligneYanc));
  ];
];

(*Écriture du fichier contenant lX(i)*X(i+j) et Y(i)*Y(i+j) avec
j∈[0:taumax]*)
Mat = SetAccuracy[mat, 3];
(*=====*)
Export[StringInsert["MatCorr.dat", NumParticule, 8], Mat, "Table"];
(*=====*)
(*      Caclul des rapport des FONCTIONS DE CORRÉLATION      *)
Cor = Table[0.0, {taumax}, {2}];

```



```

S = Table[0.0, {taumax}, {1}];
(*Écriture des entêtes*)
Cor[[1, 1]] = ToString["(*T");
Cor[[1, 2]] = ToString["Cm(t)/Cm(0)*"];

(*Écriture du premier terme de la fonction de corrélation corrigée
du 1'erreur de pointée*)
(*Cm(0) =  $\frac{1}{N} \sum_{i=1}^N X(i)^2 + Y(i)^2 - \text{ErrP}^2$  *)
S[[1]] = (0.5 * ((1 / (lgr - 1)) * Sum[mat[[i, 2]], {i, 2, lgr}]]^2
+ ((1 / (lgr - 1)) * Sum[mat[[i, 7]], {i, 2, lgr}]]^2) - ErrP * ErrP;
(*Écriture des termes Cm(0)/Cm(0)*)
Cor[[2, 2]] = S[[1]] / S[[1]];

(*Écriture des termes suivants de la fonction de corrélation *)
(*Cm(0) =  $\frac{1}{N} \sum_{i=1}^N X(i)X(i+j) + Y(i)Y(i+j)$  *)
For[ent = 3, ent < taumax + 1, ent++,
(*Écriture des temps *)
Cor[[ent, 1]] = Tex * (ent - 2);
S[[ent]] = (0.5 * ((1 / (lgr - 1)) * Sum[mat[[i, 2]], {i, 2, lgr}])) *
((1 / (lgr - 1)) * Sum[mat[[i, ent]], {i, 2, lgr}])
+ (0.5 * ((1 / (lgr - 1)) * Sum[mat[[i, (taumax + 2)]], {i, 2, lgr}])) *
((1 / (lgr - 1)) * Sum[mat[[i, (taumax + ent + 2)]], {i, 2, lgr}]);
(*Écriture des termes Cm(t)/Cm(0)*)
Cor[[ent, 2]] = S[[ent]] / S[[1]];
];

(*Écriture du fichier contenant la valeur Cm(t)/Cm(0)*)
(*=====*)
Export[StringInsert["Fcorr.dat", NumParticule, 6], Cor, "Table"];
(*=====*)

Clear["Tm", "Tm1", "Tm2", "Tm3", "modelFitTCorr1", "parsFitTCorr1",
"modelFitTCorr2", "parsFitTCorr2", "modelFitTCorr3", "parsFitTCorr3"];
(* ===== Fit de la fonctions de corrélations ===== *)
(* Suppression de l'entête du fichier *)
CorFit = Delete[Cor, {1}];

(* PREMIER Fit des données sur toutes les plages de Cm(t)/Cm(0) calculé*)
(ListPlot[{Cor, Cor}, PlotLabel -> "Extraction du  $\tau$ ||",
AxesLabel -> {"t(s)", "Cm(t)/Cm(0)"}, Joined -> {True, False},
PlotStyle -> {Blue, {Black, PointSize[0.02]}}]);
(modelFitTCorr1 =  $\left(1.0 + 2 * \left(\frac{Tm1}{Tav}\right)\right) * \left(e^{\left(-\frac{x}{Tm1}\right)}\right) - \left(2 * \left(\frac{Tm1}{Tav}\right)\right)$ ;
(* Modèle du Fit (cf publi eq 6) *)
xmax = taumax * Tac;
parsFitTCorr1 = FindFit[Select[CorFit, (0 ≤ First[#1] ≤ xmax) &],
modelFitTCorr1, {{Tm1, 0.07}}, x, MaxIterations -> 1000,
Method -> NMinimize];
(* Calcule du TEMPS DE CORRÉLATION *)
Tm = Tm1 /. parsFitTCorr1[[1]];
Tcorr1 = Tm - (Tex / 3); (* (cf publi eq 7) *)
TaumaxFit = 4 * Tm;

```

```

NbrPtFit = Ceiling[TaumaxFit / Tac];

(*OutPut1=Print[{"Premier Tm [s] :",Tm},
  {"Premier τ|| [s] :",FrameBox[Tcorr1]//DisplayForm},
  {"Premier range Fit [s] :",taumax*Tac},
  {"τ_max [ms] :",TaumaxFit},
  {"Nbr de point de Cm(t) associer à τ_max [ms] :",NbrPtFit}];*)
(*Écriture du fichier Résultat*)
TCor = Table[0.0, {4}, {2}];
TCor[[1, 1]] = ToString["(*τ||");
TCor[[1, 2]] = ToString["Δtaumax*");
TCor[[2, 1]] = Tcorr1;
TCor[[2, 2]] = xmax;

(* Deuxième Fit des données sur plages taumax=4τ|| *)

$$\left( \text{modelFitTCorr2} = \left( 1.0 + 2 * \left( \frac{Tm2}{Tav} \right) \right) * \left( e^{\left( -\frac{x}{Tm2} \right)} \right) - \left( 2 * \left( \frac{Tm2}{Tav} \right) \right); \right.$$

(* Modèle du Fit (cf publi eq 6) *)
xmax = TaumaxFit;
parsFitTCorr2 = FindFit[Select[CorFit, (0 ≤ First[#1] ≤ xmax) &],
  modelFitTCorr2, {{Tm2, 0.07}}, x, MaxIterations → 1000,
  Method → NMinimize];
(* Calcule du TEMPS DE CORRÉLATION *)
Tm = Tm2 /. parsFitTCorr2[[1]];
Tcorr2 = Tm - (Tex / 3); (* (cf publi eq 7) *)
NbrPtFit = Ceiling[(4 * Tm) / Tac];

(*OutPut2=Print[{"Deuxième Tm [s] :",Tm},
  {"Deuxième τ|| [s] :",FrameBox[Tcorr2]//DisplayForm},
  {"Deuxième range Fit [s] :",xmax},
  {"τ_max [ms] :",4*Tm},
  {"Nbr de point de Cm(t) associer à τ_max [ms] :",NbrPtFit}];*)
TCor[[3, 1]] = Tcorr2;
TCor[[3, 2]] = xmax;

(* Troisième Fit des données sur plages taumax=4τ|| *)

$$\left( \text{modelFitTCorr3} = \left( 1.0 + 2 * \left( \frac{Tm3}{Tav} \right) \right) * \left( e^{\left( -\frac{x}{Tm3} \right)} \right) - \left( 2 * \left( \frac{Tm3}{Tav} \right) \right); \right.$$

(* Modèle du Fit (cf publi eq 6) *)
xmax = 4 * Tm;
parsFitTCorr3 = FindFit[Select[CorFit, (0 ≤ First[#1] ≤ xmax) &],
  modelFitTCorr3, {{Tm3, 0.07}}, x, MaxIterations → 1000,
  Method → NMinimize];
(* Calcule du TEMPS DE CORRÉLATION *)
Tm = Tm3 /. parsFitTCorr3[[1]];
Tcorr3 = (Tm - (Tex / 3)); (* (cf publi eq 7) *)
NbrPtFit = Ceiling[(4 * Tm) / Tac];

(*OutPut3=Print[{"Troisième Tm [s] :",Tm},
  {"Troisième τ|| [s] :",FrameBox[Tcorr3]//DisplayForm},
  {"Troisième range Fit [s] :",xmax},
  {"τ_max [ms] :",4*Tm},
  {"Nbr de point de Cm(t) associer à τ_max [ms] :",NbrPtFit}];*)

```

```

(*Écriture du fichier Résultat*)
TCor[[4, 1]] = Tcorr3;
TCor[[4, 2]] = xmax;
InfoFile[[j, 3]] = Tcorr3;
InfoFile[[j, 4]] = xmax;
TpsCorrelation[[j, 1]] = Tcorr3;
TpsTaumax[[j, 1]] = xmax;
(* ===== Calcul de Aeq ===== *)
Clear["AeqRaw", "Acorr", "écritureRexp"];
AeqRaw = Table[table[[i, NligneAeq]], {i, nCSV, Dim[[1]]}];
Acorr = Table[0.0, {(lgr)}];

For[i = (1), i < ((lgr + 1)), i++,
  Acorr[[i]] =
    AeqRaw[[i]] /
      (((2 * Tcorr3) / Tex) - 2 * (((Tcorr3 / Tex) ^ 2) * (1 - Exp[- (Tex / Tcorr3)]))) ^
        (1 / 2));
];
(* =====Ecriture dans histoRexp.dat===== *)
Acut = SetAccuracy[Acorr, 4];
écritureRexp = PrependTo[Acut, NumParticule];
écritureRexp = ArrayReshape[écritureRexp, {1, nPtFichierCSV}, 0.];
écritureRexp = Flatten[écritureRexp];
EcritureRexp = Table[{EcritureRexp[[i]], écritureRexp[[i]]},
  {i, 1, (nPtFichierCSV)}];
EcritureRexp = Flatten[EcritureRexp];
EcritureRexp = Partition[EcritureRexp, (cptFileUsed + 1)];
(* =====
  Calcul de Aeq Corriger  $\tau$ ||
  puis MOY et écriture dans InfoFile ===== *)
MeanAeqCorr[[j]] = Mean[Acorr];
(* ===== Calcul de Aeq MOY et écriture dans InfoFile ===== *)
AeqMoy = Part[DataMoy2[[j - 2]]];
(* AeqMoyCorrige : AeqMoy pour chaque trajectoire corriger par  $\tau$ ||*)
AeqMoyTcor =
  ((Part[DataMoy2[[j - 2]]]) /
    (((2 * Tcorr3) / Tex) - 2 * (((Tcorr3 / Tex) ^ 2) * (1 - Exp[- (Tex / Tcorr3)]))) ^
      (1 / 2)));
InfoFile[[j, 5]] = AeqMoy;
InfoFile[[j, 6]] = AeqMoyTcor;
InfoFile[[j, 7]] = MeanAeqCorr[[j]]; (*J?*)
histoAeqMoyClean[[j, 1]] = AeqMoy;
histoAeqMoyCleanTcor[[j, 1]] = AeqMoyTcor;
InfoFile = SetAccuracy[InfoFile, 5];
ResultCor = SetAccuracy[Join[Cor, TCor], 8];
(*=====*)
Export[StringInsert["Fcorr.dat", NumParticule, 6], ResultCor, "Table"];
Export["InfoFile.dat", InfoFile, "Table"];
(*=====*)
];

```

Écriture des fichiers **R** exp|| corrigées de  $\tau_{||}$  :

```
(* ====Ecriture dans histoRexp.dat===== *)
Export["HistoRexp.dat", EcritureRexp, "Table"];
(* RE criture de l'histogramme des AeqMOY avec correction du a  $\tau$  | *)
AA = Join [Transpose[histoAeqMoyClean], Transpose[histoAeqMoyCleanTcor]];
BB = Transpose[AA];
CC = SetAccuracy[BB, 5];
(*=====*)
NomHistoMoy3 = ToString[StringForm["histoAeqMoySup`Tcorr.dat", (AmpMin)]];
Export[NomHistoMoy3, CC, "Table"];
(*=====*)
```

## Repr sentation/histo

```
NomHisto3 = ToString["HistoRexp.dat"];
(* Delete des donn es des entetes pour tracer les diagrammes*)
Data3 = Import[NomHisto3, "Table"];
Data3 = Delete[Data3, {1}];
Data3 = Transpose[Data3];
Data3 = Delete[Data3, {1}];
data3 = Flatten[Data3];
data3 = DeleteCases[data3, 0.];
lgrhisto3 = Length[data3];

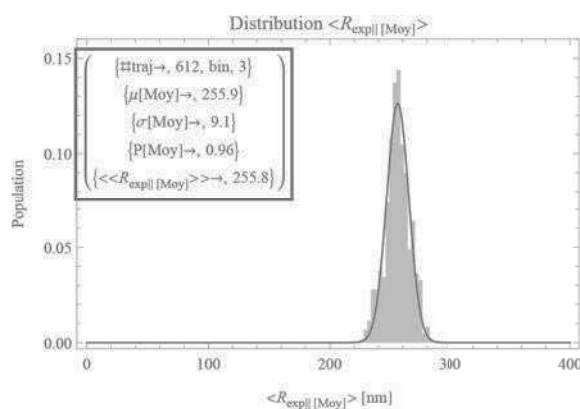
NomHistoMoy3 = ToString[StringForm["histoAeqMoySup`Tcorr.dat", (AmpMin)]];
(* Delete des donn es des entetes pour tracer les diagrammes*)
DataMoy3 = Import[NomHistoMoy3, "Table"];
DataMoy3 = DataMoy3[[All, 2]];
dataMoy3 = Delete[DataMoy3, {{1}, {2}}];
lgrhistoMoy3 = Length[dataMoy3];
```

Représentation/histo  $\langle R_{\text{exp}} \rangle$  corrigées de  $\tau_{||}$ :

```
(* Amplitudes Corrigée par  $\tau_{||}$  → AeqCleanTcor *)
Histogram[dataMoy3, {Amin, Amax, sb2}, "Probability"];
{bins, counts} = HistogramList[dataMoy3, {Amin, Amax, sb2}, "Probability"];
centers = MovingAverage[bins, 2];

modelMoy3 = sb2 * 
$$\left( P1 \right) * \left( \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi} \sigma} \right);$$


parsMoy3 = FindFit[{centers, counts}^T, modelMoy3,
  {{ $\mu$ , 200}, { $\sigma$ , 10}, {P1, 0.1}}, x];
H5 = Show[Histogram[dataMoy3, {Amin, Amax, sb2}, "Probability",
  PlotLabel → "Distribution  $\langle R_{\text{exp}} \rangle$  [Moy]", Frame → True,
  FrameLabel → {" $\langle R_{\text{exp}} \rangle$  [Moy]", "Population"},
  ChartLabels →
  {Placed[
    Framed[MatrixForm[{{{"#traj" →, lgrhistoMoy3, "bin", sb2},
      {" $\mu$  [Moy] →", SetAccuracy[ $\mu$  /. parsMoy3[[1]], 2}},
      {" $\sigma$  [Moy] →", SetAccuracy[ $\sigma$  /. parsMoy3[[2]], 2}},
      {"P [Moy] →", SetAccuracy[P1 /. parsMoy3[[3]], 3}},
      {" $\langle R_{\text{exp}} \rangle$  [Moy] →", SetAccuracy[Mean[dataMoy3], 2]}}]], {0.2, 0.8}]],
  ChartStyle → Gray], Plot[modelMoy3 /. parsMoy3 // Evaluate,
  {x, Amin, Amax}, PlotRange → All, PlotStyle → Black]]
```



Représentation/histo  $R_{\text{exp}}$  corrigées de  $\tau$ :

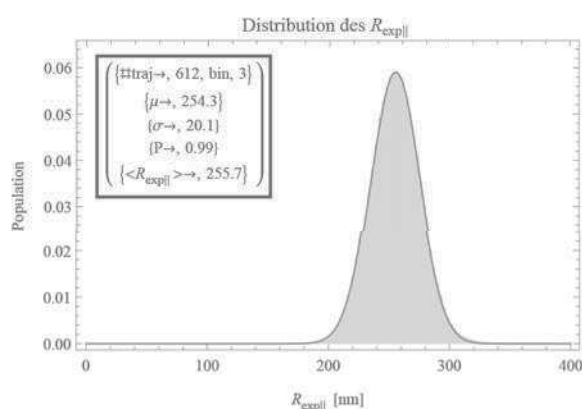
```
(
Histogram[data3, {Amin, Amax, sb2}, "Probability"];

{bins, counts} = HistogramList[data3, {Amin, Amax, sb2}, "Probability"];
centers = MovingAverage[bins, 2];

model3 = sb2 * (P1) *  $\left( \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \right)$ ;

pars3 = FindFit[{centers, counts}^T, model3, {{μ, 200}, {σ, 8}, {P1, 0.5}}, x];
H4 = Show[Histogram[data3, {Amin, Amax, sb2}, "Probability",
  PlotLabel → "Distribution des  $R_{\text{exp}}$ ", Frame → True,
  FrameLabel → {" $R_{\text{exp}}$  [nm]", "Population"},
  ChartLabels →
    {Placed[Framed[MatrixForm[{{"#traj→", lgrhistoMoy3, "bin", sb2},
      {"μ→", SetAccuracy[μ /. pars3[[1]], 2]},
      {"σ→", SetAccuracy[σ /. pars3[[2]], 2]},
      {"P→", SetAccuracy[P1 /. pars3[[3]], 3]},
      {"< $R_{\text{exp}}$ >→", SetAccuracy[Mean[data3], 2]}]}], {0.2, 0.8}]}],
  Plot[model3 /. pars3 // Evaluate, {x, Amin, Amax}, PlotRange → All]]

(* REÉcriture des données globales *)
Donnee[[4, 1]] = ToString["AeqCleanTcor"];
Donnee[[4, 2]] = (μ /. pars3[[1]]);
Donnee[[4, 3]] = ToString["AeqCleanTcorMoy"];
Donnee[[4, 4]] = (μ /. parsMoy3[[1]]);
Donnee[[5, 1]] = ToString["Mean[TpsCorrelation]"];
Donnee[[5, 2]] = Mean[Flatten[TpsCorrelation]];
Donnee[[5, 3]] = ToString["Mean[TpsTaumax]"];
Donnee[[5, 4]] = Mean[Flatten[TpsTaumax]];
donnee = SetAccuracy[Donnee, 6];
(*=====*)
Export["DataGlobal.dat", donnee, "Table"];
(*=====*)
```



SSi 2 populations -- Représentation de l'ensemble des points d'amplitude

-- TRAITEMENT DES  $\tau_{||}$  --

Distribution  $\langle \tau_{||} \rangle$  (TEMPS DE CORRÉLATION MOYENNÉ SUR TOUTE LA DURÉE DE LA TRAJECTOIRE)

Exportation des  $\langle \tau_{||} \rangle$  (TEMPS DE CORRÉLATION MOYENNÉ SUR TOUTE LA DURÉE DE LA TRAJECTOIRE)

```
DataInfoFile = Import["InfoFile.dat", "Table"];
```

Représentation/histo  $\langle \tau_{||} \rangle$  :

```
dataInfoFile = Delete[DataInfoFile, {{1}, {2}}];
dataTau = (dataInfoFile[[All, 3]] * 1000) (*ms*);
Print["# Traj = ", lgrTau = Length[dataTau]];

(* Variable de construction des histogrammes*)
Tmax = Max[dataTau] + 20;
If[Tmax < 30,
  Tmax = TauCaracADN + 20;
  Print[Style["== ATTENTION ==", 15, Red]];
  Print["Echelle axe mis à [0: ", Tmax, "] [ms]"];
  Tmax = TauCaracADN + 20;
]

Tmin = 0;
sb3 = 3; (*valeur moyenne*)
(* 1er calcul pour definir la taille des bins*)
(
Histogram[dataTau, {Tmin, Tmax, sb3}, "Probability"];

{bins, counts} = HistogramList[dataTau, {Tmin, Tmax, sb3}, "Probability"];
centers = MovingAverage[bins, 2];

modelTau = sb3 * (
  (P1) * 
$$\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}$$

);

parsTau = FindFit[{centers, counts}^T, modelTau, {{μ, 25}, {σ, 5}, {P1, 0.5}},
  x];

(*Print["μ[τ||] (FitG) = ", *) MTau = μ /. parsTau[[1]] (*, " ms"*));
sigmaTau = σ /. parsTau[[2]];
(*Print["Bornes Sup = ", *) TauMaxbis = Ceiling[MTau + 2.5 * sigmaTau]
(*, " ms"*));
(*Print["Bornes Inf = ", *) TauMinbis = Floor[MTau - 4 * sigmaTau] (*, " ms"*));

HDtau = Show[Histogram[dataTau, {Tmin, Tmax, sb3}, "Probability",
  PlotLabel → "Distribution τ||", Frame → True,
  FrameLabel → {"Tps de correlation [ms]", "Population"},
  ChartLabels →
    {Placed[Framed[MatrixForm[{"#traj -> ", lgrTau},
```

```

      {"μ -> ", SetAccuracy[μ /. parsTau[[1]], 2]},
      {"σ -> ", SetAccuracy[σ /. parsTau[[2]], 2]},
      {"P1-> ", SetAccuracy[P1 /. parsTau[[3]], 3]},
      {"<τ||> -> ", SetAccuracy[Mean[dataTau], 2]},
      {"+/-2.5σ -> [", TauMinbis, ":", TauMaxbis, "]"}}]], {0.7, 0.8}]],
ChartStyle -> Orange], Plot[modelTau /. parsTau // Evaluate,

{x, Tmin, Tmax}, PlotRange -> All, PlotStyle -> Black]]
];

sb3 = Ceiling[(TauMaxbis - TauMinbis) / Sqrt[lgrTau]];
If[sb3 > 3,
  Print[Style["== ATTENTION ==", 15, Red]];
  Print["Taille du bin = ", sb3 = Ceiling[(TauMaxbis - TauMinbis) / Sqrt[lgrTau]],
    "[ms]"];
  Print[Style["MAIS bins mit à 3[ms]", 13, Red]];
  sb3 = 3;
]

(
Histogram[dataTau, {Tmin, Tmax, sb3}, "Probability"];

{bins, counts} = HistogramList[dataTau, {Tmin, Tmax, sb3}, "Probability"];
centers = MovingAverage[bins, 2];

modelTau = sb3 * (P1) * 
$$\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}$$
;

parsTau = FindFit[{centers, counts}^T, modelTau, {{μ, 25}, {σ, 5}, {P1, 0.5}},
  x];

Print["μ[τ||] (FitG) = ", MTau = μ /. parsTau[[1]], " ms"];
sigmaTau = σ /. parsTau[[2]];
Print["Bornes Sup = ", TauMaxbis = Ceiling[MTau + 2.5 * sigmaTau], " ms"];
Print["Bornes Inf = ", TauMinbis = Floor[MTau - 2.5 * sigmaTau], " ms"];

If[TauMaxbis < (TauCaracADN),
  Print[Style["== ATTENTION ==", 15, Red]];
  Print[Style["Borne Sup=TauMax < Tau Caracteristique de l'ADN utilisé : ",
    13, Red]];
  Print["Tau Carac ADN : ~", TauCaracADN, "[ms]"];
  Print[Style["MAIS Borne Sup mit à ", 13, Red]];
  Print["TauMax : ", TauMaxbis = Ceiling[TauCaracADN + 2.5 * sigmaTau], " [ms]"];
];

If[TauMinbis < (13.0),
  Print[Style["== ATTENTION ==", 15, Red]];
  Print[Style["Borne Inf < Tau corrigable : ", 13, Red]];
  Print[Style["DONC Borne INF mit à ", 13, Red]];
  Print["TauMin : ", TauMinbis = 13.0, " [ms]"];
];

```



```

HDtau = Show[Histogram[dataTau, {Tmin, Tmax, sb3}, "Probability",
  PlotLabel -> "Distribution  $\tau$ ||", Frame -> True,
  FrameLabel -> {"Tps de correlation [ms]", "Population"},
  ChartLabels ->
    {Placed[Framed[MatrixForm[{{{"#traj" -> " , lgrTau , "bin", sb3},
      {" $\mu$  -> " , SetAccuracy[ $\mu$  /. parsTau[[1]], 2}},
      {" $\sigma$  -> " , SetAccuracy[ $\sigma$  /. parsTau[[2]], 2}},
      {"P1-> " , SetAccuracy[P1 /. parsTau[[3]], 3}},
      {"< $\tau$ > -> " , SetAccuracy[Mean[dataTau], 2}},
      {"-5,+2.5] $\sigma$  -> [ , TauMinbis, ":", TauMaxbis, "]"}}}]],
      {0.75, 0.8}]], ChartStyle -> Orange],
  Plot[modelTau /. parsTau // Evaluate, {x, Tmin, Tmax}, PlotRange -> All,
    PlotStyle -> Black]]

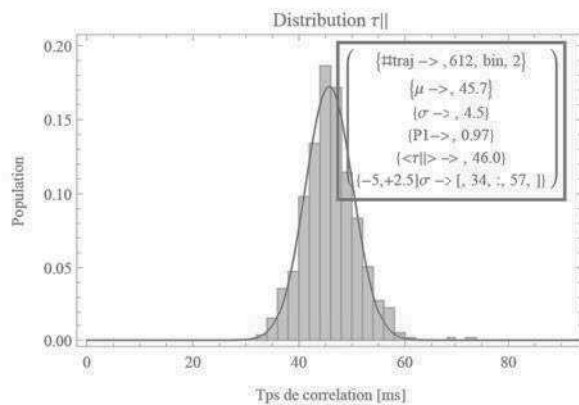
```

# Traj = 612

$\mu[\tau]$  (FitG) = 45.71 ms

Bornes Sup = 57 ms

Bornes Inf = 34 ms



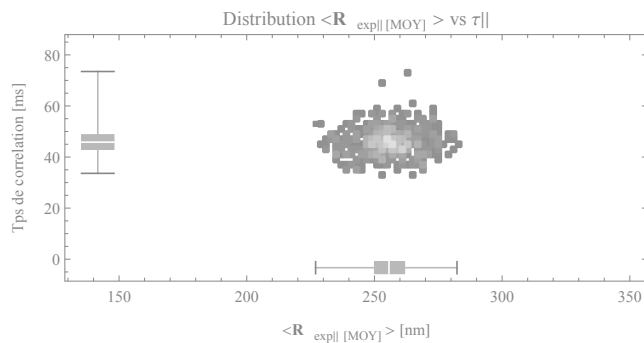
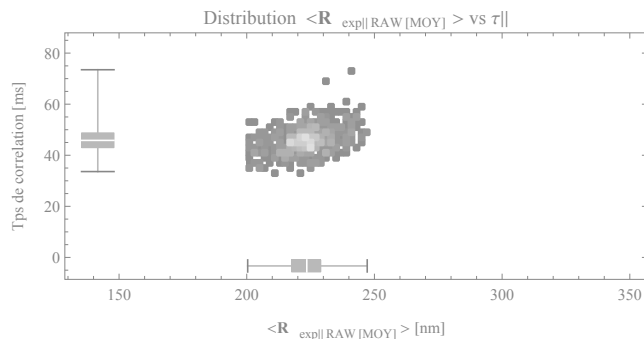
SSI 2 populations -- Représentation de l'ensemble des points d'amplitude

Représentation/Matrice de densité des  $\langle \tau_{||} \rangle$  :

```
DataMdTauMoy = Table[{dataInfoFile[[i, 5]], dataInfoFile[[i, 3]] * 1000},
  {i, 1, lgrTau}] ; (*ms*)
DataMdTau = Table[{dataInfoFile[[i, 6]], dataInfoFile[[i, 3]] * 1000},
  {i, 1, lgrTau}] ; (*ms*)

davantcorr = DensityHistogram[DataMdTauMoy, {{150, 350, 2}, {00, 80, 2}},
  AspectRatio → Automatic, PlotLabel → "Distribution  $\langle R_{exp} || RAW [MOY] \rangle$  vs  $\tau ||$ ",
  FrameLabel → {" $\langle R_{exp} || RAW [MOY] \rangle$  [nm]", "Tps de correlation [ms]"},
  ColorFunction → "SolarColors", Method → {"DistributionAxes" → "BoxWhisker"}]

daprescorr = DensityHistogram[DataMdTau, {{150, 350, 2}, {00, 80, 2}},
  AspectRatio → Automatic, PlotLabel → "Distribution  $\langle R_{exp} || [MOY] \rangle$  vs  $\tau ||$ ",
  FrameLabel → {" $\langle R_{exp} || [MOY] \rangle$  [nm]", "Tps de correlation [ms]"},
  ColorFunction → "SolarColors", Method → {"DistributionAxes" → "BoxWhisker"}]
```



4eme FILTRE de Sélection :

-> Filtre sur  $\langle \tau_{||} \rangle$  (TEMPS DE CORRÉLATION MOYENNÉ SUR TOUTE LA DURÉE DE LA TRAJECTOIRE)

Extraction des numéro des trajectoires Ok&Fausse

```
Dataglobale = Table[{dataInfoFile[[i, 2]], dataInfoFile[[i, 3]] * 1000},
  {i, 1, lgrTau}];

Numapparition = Table[0.0, {lgrTau}, {1}];
NumTrajOK = Table[0.0, {lgrTau}, {1}];
NumTrajNON = Table[0.0, {lgrTau}, {1}];
Matrixcpt = Table[0.0, {lgrTau}, {1}];
IndiceTrajNON = Table[0.0, {lgrTau}, {1}];
cptOK = 0;
cptNON = 0;

For[i = 1, i < (lgrTau + 1), i++,
  If [Dataglobale[[i, 2]] > TauMinbis && Dataglobale[[i, 2]] < TauMaxbis,
    (cptOK = cptOK + 1;
    NumTrajOK[cptOK] = Dataglobale[[i, 1]];
    Matrixcpt[[i]] = 1;
    Numapparition[[i]] = Dataglobale[[i, 1]]);
    (cptNON = cptNON + 1;
    NumTrajNON[cptNON] = Dataglobale[[i, 1]];
    Matrixcpt[[i]] = 0;
    IndiceTrajNON[cptNON] = i;
    Numapparition[[i]] = Dataglobale[[i, 1]]);
  ]
];
Print["Vérification : #_traj[Tot]= ", lgrTau, " => #_traj[Ok]= ",
  cptOK, " + #_traj[No]= ", cptNON]
NumTrajOK = Table[{NumTrajOK[[i]]}, {i, 1, cptOK}];
NumTrajNON = Table[{NumTrajNON[[i]]}, {i, 1, cptNON}];
IndiceTrajNON = Table[{IndiceTrajNON[[i]]}, {i, 1, cptNON}];
iindice = Table[i, {i, 1, lgrTau}];
```

Vérification : #\_traj[Tot]= 612 => #\_traj[Ok]= 600 + #\_traj[No]= 12

Écriture dans le fichier InfoFile

```
For[i = 1, i < (lgrTau + 1), i++,
  DataInfoFile[[i + 2, 8]] = Numapparition[[i]];
  DataInfoFile[[i + 2, 9]] = iindice[[i]];
  DataInfoFile[[i + 2, 10]] = Matrixcpt[[i]];
]
Export["InfoFile.dat", DataInfoFile, "Table"]
InfoFile.dat
```

Verification du nombre de fichier .csv AVANT suppression en fonction de Tau

```
Fichiertau = FileNames["*.csv"];
Print["# de fichier : ", DimFtau = Dimensions[Fichiertau]];
# de fichier : {612}
```

```

DataRexp = Import["HistoRexp.dat", "Table"];

dataRexp = Delete[DataRexp, {{1}}]; (*Suppresion N° traj*)
dataRexp = Transpose[dataRexp];
dataRexp = Delete[dataRexp, {{1}}]; (*Suppresion colonne String[Rexp]*)
dataRexp = Transpose[dataRexp];

For[i = 1, i < (cptNON + 1), i++,
  indicej = IndiceTrajNON[[i, 1]];
  If[Numapparition[[indicej]] == NumTrajNON[[i, 1]],
    (NomFaSuppr1 = ToString[StringForm["a_a_a_ResultsParticle` `.csv",
      (NumTrajNON[[i, 1]])]]);
    NomFaSuppr2 = ToString[StringForm["MatCorr` `.dat", (NumTrajNON[[i, 1]])]]];
    NomFaSuppr3 = ToString[StringForm["Fcorr` `.dat", (NumTrajNON[[i, 1]])]]];

    NomFaSupprrenomme =
      ToString[StringForm["T_a_a_a_ResultsParticle` `.csv",
        (NumTrajNON[[i, 1]])]]];
    (*Sauvegarde du fichier + renommé T_...*)
    FaSuppr = Import[NomFaSuppr1, "Table"];
    NomArbosave = ToString[StringForm["` `/saves/", (directory)]];
    SetDirectory[NomArbosave];
    Export[NomFaSupprrenomme, FaSuppr, "Table"];
    SetDirectory["../"];

    NomArbocourant1 = ToFileName[directory, NomFaSuppr1];
    NomArbocourant2 = ToFileName[directory, NomFaSuppr2];
    NomArbocourant3 = ToFileName[directory, NomFaSuppr3];

    Print["A supprimer n°Traj : ", NumTrajNON[[i, 1]]];
    (*Suppression du fichier a_a...*)
    cmdsuppr1 = ToString[StringForm["rm ` `.", (NomArbocourant1)]];
    cmdsuppr2 = ToString[StringForm["rm ` `.", (NomArbocourant2)]];
    cmdsuppr3 = ToString[StringForm["rm ` `.", (NomArbocourant3)]];
    Run[ToString[cmdsuppr1], ToString[cmdsuppr2], ToString[cmdsuppr3]];
    (*DeleteFile[ToFileName[directory, ToString[NomFaSupprrenomme]]]*)

    (*Suppression de la colonne correspondante dans HistoRexp.dat*)
    dataRexp = Transpose[dataRexp];
    dataRexp = Delete[dataRexp, {{IndiceTrajNON[[i, 1]] - i + 1}}];
    dataRexp = Transpose[dataRexp];
  ),
  (
    Print["Attention : Erreur "]
  )
];

Fichierta2 = FileNames["*.csv"];
Print["# de fichier : ", DimFta2 = Dimensions[Fichierta2]];

Export["HistoRexp_Final.dat", dataRexp, "Table"];

```

```
A supprimer n°Traj : 1 000 229
A supprimer n°Traj : 1 000 290
A supprimer n°Traj : 1 000 340
A supprimer n°Traj : 1 000 346
A supprimer n°Traj : 1 000 395
A supprimer n°Traj : 1 000 400
A supprimer n°Traj : 1 000 431
A supprimer n°Traj : 1 000 632
A supprimer n°Traj : 1 000 639
A supprimer n°Traj : 100 078
A supprimer n°Traj : 2 000 443
A supprimer n°Traj : 3 000 364
# de fichier : {600}
```

Verification du nombre de fichier .csv APRÈS suppression en fonction de Tau

```
Fichiertau2 = FileNames["*.csv"];
Print["# de fichier : ", DimFtau2 = Dimensions[Fichiertau2]];
# de fichier : {600}
```

Représentation des **R<sub>exp||</sub>** final

```

Rexp = Import["HistoRexp_Final.dat", "Table"];
(*Supprimer les Zero !***)
```

$$\text{model4} = \text{sb4} * \left( (\text{P1}) * \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \right);$$

```

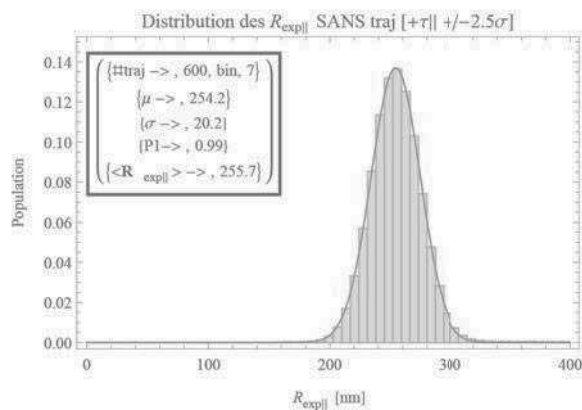
Rexp = Delete[Rexp, {1}];
Rexp = Flatten[Rexp];
Rexp = DeleteCases[Rexp, 0.];
LRexp = Length[Rexp];
Dimensions[Rexp];

Rm = ( $\mu$  /. pars3[[1]]);
Rs = ( $\sigma$  /. pars3[[2]]);
RMin = SetAccuracy[Rm - 4 * Rs, 2];
RMax = SetAccuracy[Rm + 4 * Rs, 2];
Amax = 400;
Amin = 0;
sb4 = Ceiling[(RMax - RMin) / Sqrt[(DimFtau2[[1]] + 50)]];
If[sb4 > 15,
  Print[Style["== ATTENTION ==", 15, Red]];
  Print["Taille du bin = ", sb4 "[ms]"];
  Print[Style["MAIS bins mit à 20[nm]", 13, Red]];
  sb4 = 20;
]

Histogram[Rexp, {Amin, Amax, sb4}, "Probability"];
{bins, counts} = HistogramList[Rexp, {Amin, Amax, sb4}, "Probability"];
centers = MovingAverage[bins, 2];

(pars4 = FindFit[{centers, counts}^T, model4, {{ $\mu$ , 200}, { $\sigma$ , 10}, {P1, 0.5}}, x];
HRExp = Show[Histogram[Rexp, {Amin, Amax, sb4}, "Probability",
  PlotLabel → "Distribution des Rexp|| SANS traj [+ $\tau$ ] +/-2.5 $\sigma$ ",
  Frame → True, FrameLabel → {"Rexp|| [nm]", "Population"},
  ChartLabels →
    {Placed[
      Framed[MatrixForm[{{"#traj -> ", DimFtau2[[1]], "bin", sb4 },
        {" $\mu$  -> ", SetAccuracy[ $\mu$  /. pars4[[1]], 2}},
        {" $\sigma$  -> ", SetAccuracy[ $\sigma$  /. pars4[[2]], 2}},
        {"P1-> ", SetAccuracy[P1 /. pars4[[3]], 3}},
        {"<Rexp|| > -> ", SetAccuracy[Mean[Rexp], 2]}]]], {0.2, 0.8}]]],
  Plot[model4 /. pars4 // Evaluate, {x, Amin, Amax}, PlotRange → All]]

```



"Sous-Compactage" des (Xcor,Ycor) pour un "champ" de manip

Initialisation compteur

```
cptMXY = 0;
```

"sous-compactage"

```
OutputName = ToString[StringJoin[{"XYposition_", ParametreManip}]];
(* ===== Écriture du fichier GLOBAL des résultats ===== *)
FichierXY = FileNames["MatCorr*.dat"];
Print["Vérife : # Traj =", DimFMXY = Dimensions[FichierXY]];
NbrCXcor = 1;
NbrCYcor = 14;

For[j = 1, j < (DimFMXY[[1]] + 1), j++,
  cptMXY = cptMXY + 1;
  NomArbo = ToFileName[directory, FichierXY[[j]]];
  data = Import[NomArbo, "Table"];
  data = Delete[data, {1}];
  data = Table[{data[[i]][[NbrCXcor]], data[[i]][[NbrCYcor]]},
    {i, 1, Length[data]}];
  data = SetAccuracy[data, 4];

  file = OpenAppend[OutputName];
  WriteString[file, ExportString[data, "Table"], "\n"];
  Close[file];

];

(*Vérification si toutes les trajectoires ont été prise en compte*)
Print["# traj 'compacté' = ", cptMXY]
```

```
Vérife : # Traj = {600}
```

```
# traj 'compacté' = 600
```

## Représentation

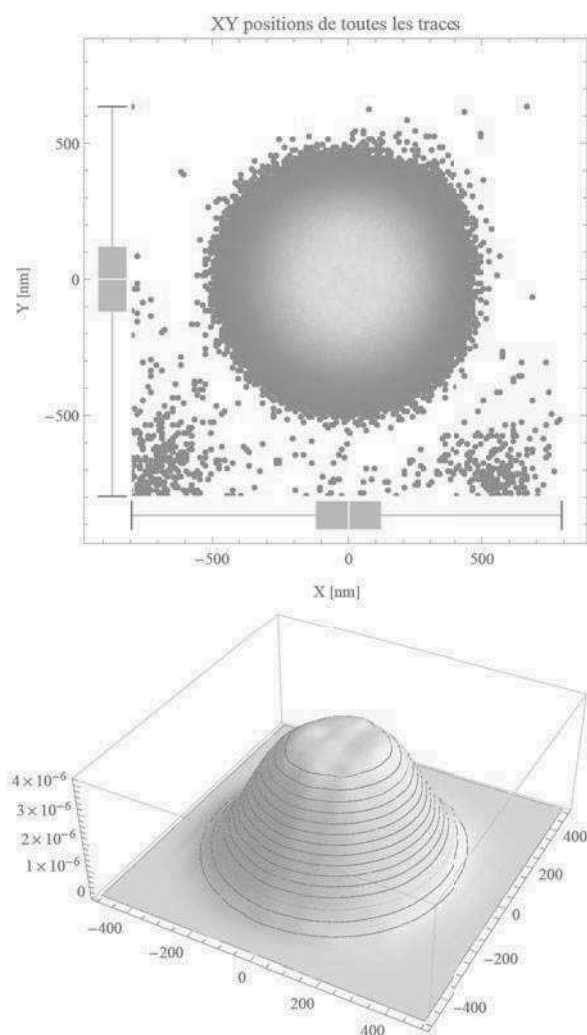
```

NomArboTot = ToFileName[directory, OutputName];
dataTotXY = Import[NomArboTot, "Table"];
dataTotXY = Table[{dataTotXY[[i]][[1]], dataTotXY[[i]][[2]]},
  {i, 1, Length[dataTotXY]}];

pdensity1 = DensityHistogram[dataTotXY, {{-800, 800, 10}, {-800, 800, 10}},
  AspectRatio → Automatic, PlotLabel → "XY positions de toutes les traces",
  FrameLabel → {"X [nm]", "Y [nm]"}, ColorFunction → "SolarColors",
  Method → {"DistributionAxes" → "BoxWhisker"}]

pdensity2 =
  Show[SmoothHistogram3D[dataTotXY, PlotRange → {{-500, 500}, {-500, 500}}]]

```





Suppression des points apres Compilation de toutes les données

**Pour la DÉNATURATION nous n'utiliserons pas ce filtre sauf si une population de bille collées qui est significativement visible et séparer de la population principale. Cependant pour ne pas modifier tout le programme nous lançons cette étapes avec des bornes non-sélectives**

Définition des bornes de suppression de points

```
AmpMaxbis = Ceiling[( $\mu$  /. pars4[[1]]) + 2.5 * ( $\sigma$  /. pars4[[2]])];
(*AmpMaxbis=Ceiling[( $\mu$ 2/.pars4[[4]])+2.5*( $\sigma$ /.pars4[[2]])];*)
AmpMinbis = Floor[( $\mu$  /. pars4[[1]]) - 2.5 * ( $\sigma$  /. pars4[[2]])];

Print["Ssi  $R \in \mu[R_{exp}] \pm 5\sigma$  : [", AmpMinbis, " : ", AmpMaxbis, "]"]
```

Ssi  $R \in \mu[R_{exp}] \pm 5\sigma$  : [203 : 305]

Suppression des points et nouveau fichier.dat

```
databis = Table[0.0, {Length[Rexp]}, {1}];
cpt = 0;
cpttrash = 0;
For[i = 1, i < (Length[Rexp]), i++
  If [Rexp[[i]] > AmpMinbis && Rexp[[i]] < AmpMaxbis,
    (cpt = cpt + 1;
     databis[[cpt]] = Rexp[[i]]);
    , (cpttrash = cpttrash + 1;
     )
  ]
];
Print["#Tot: ", Length[Rexp], " #Gardé: ", cpt, " #Trash: ", cpttrash,
      " Vérification (A+B): ", cpt+cpttrash]
(*=====*)
NomDataNEW = ToString[StringForm["HistoRexp_Final_NEW`.dat", (AmpMinbis)]];
databis = Flatten[databis];
databis = DeleteCases[databis, 0.0];
Print["New_#Points: ", Length[databis]]
Export[NomDataNEW, databis, "Table"];
(*=====*)
```

#Tot: 1 734 730 #Gardé: 1 693 797 #Trash: 40 932 Vérification (A+B): 1 734 729

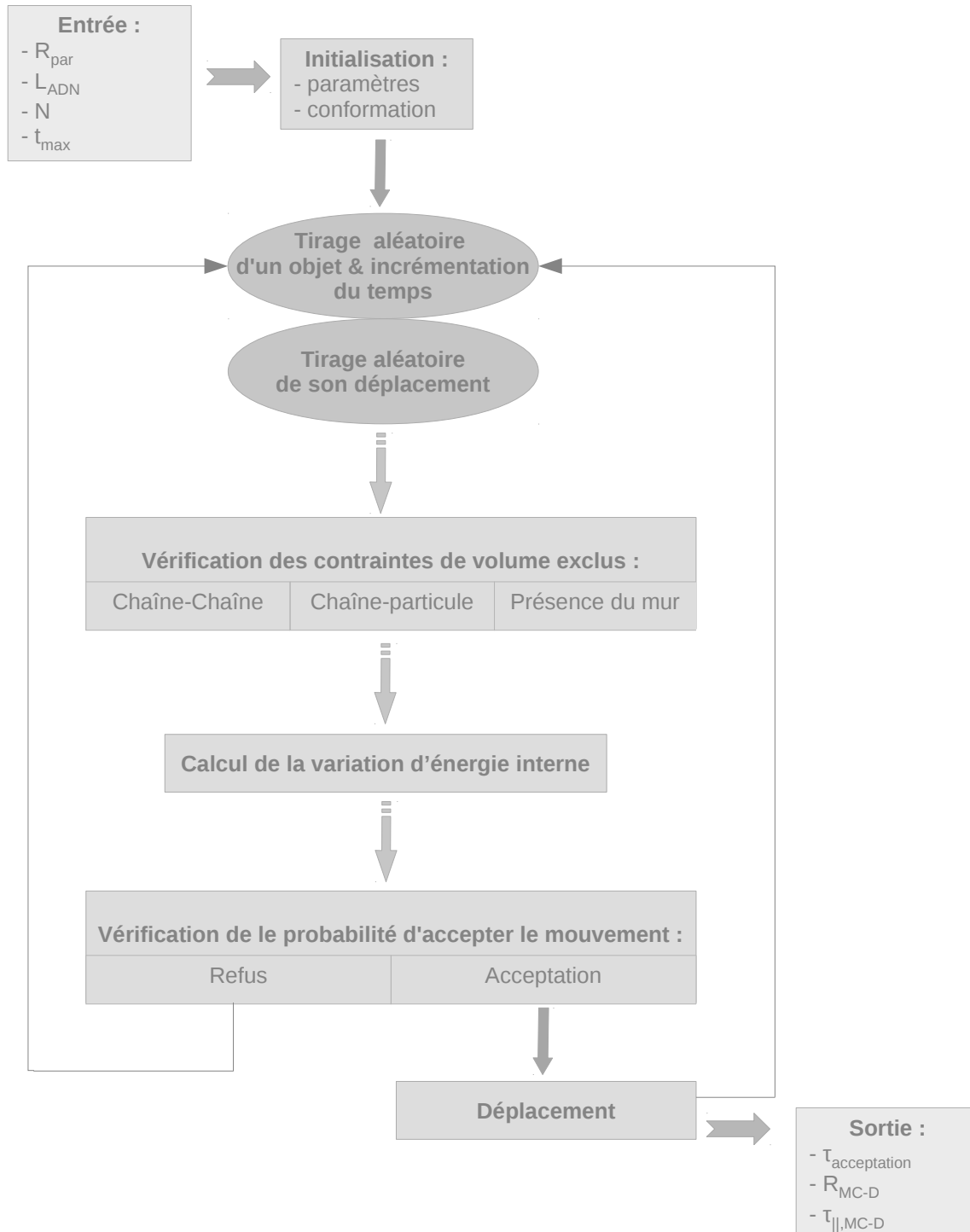
New\_#Points: 1 693 797

Exportation du nouveau fichier.dat

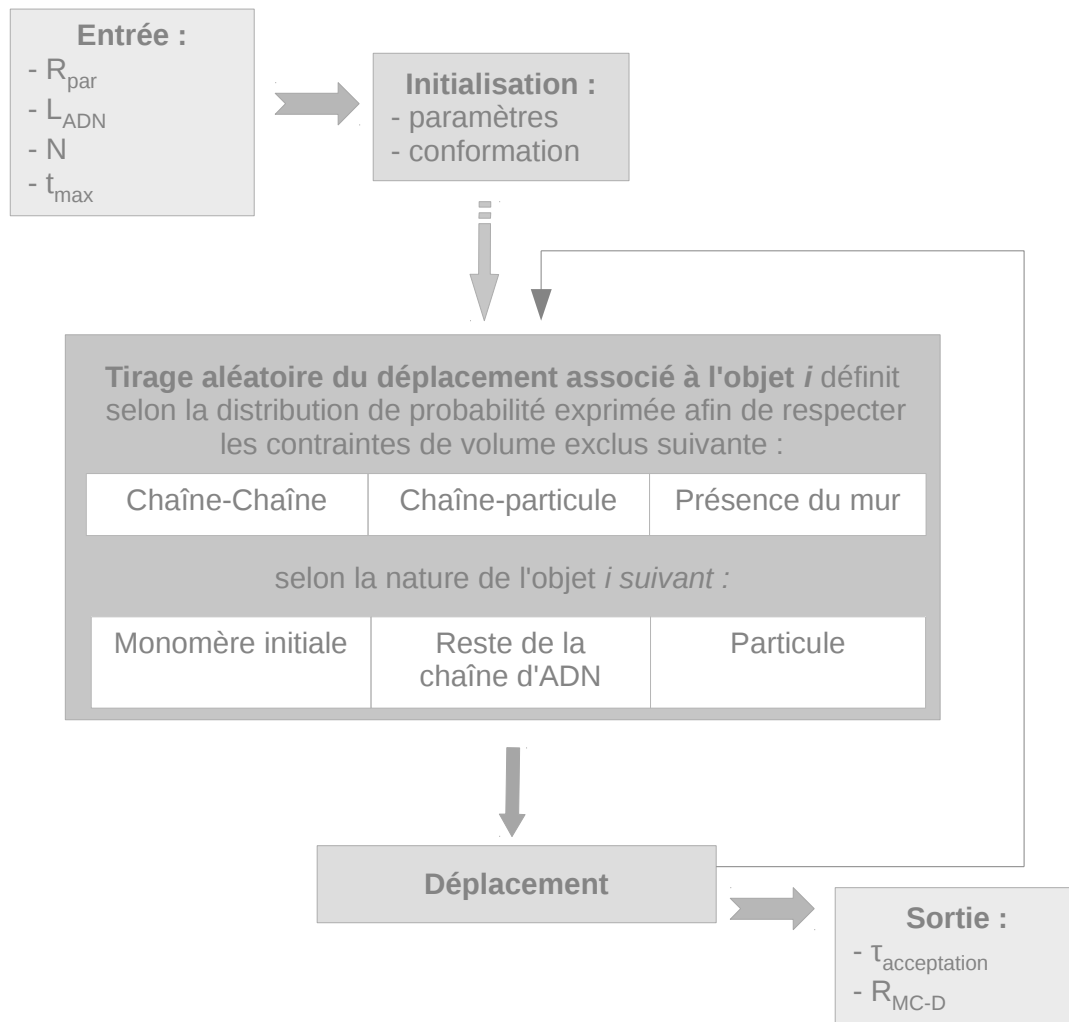
```
DataNew = Import[NomDataNEW, "Table"];
lgrhistoNew = Length[DataNew];
DataNew = DataNew[[All, 1]];
```

## D SCRIPT DES PROGRAMMES NUMÉRIQUES

### D.1 Logigramme du modèle Monte-Carlo Dynamique (MC-D)



## D.2 Logigramme du modèle Échantillonnage Exacte (EE)



## E DONNÉES GÉNÉRALES

### E.1 Données TPM, courbe de calibration

$L_{ADN}$ bp	$A_{eq}$ $\boxed{2D}$ nm	$R_{  }$ $\boxed{2D}$ nm	$\tau_{  }$ ms	$R_{ADN}$ $\boxed{3D}$ nm	$R_{ADN,S}$ $\boxed{3D}$ nm	$L_p$ (FJC) nm	$L_p$ (FJC,S) nm	$L_p$ (WLC) nm	$L_p$ (WLC,S) nm
583	112	151	20	108	126	29	40	35	45
639	119	157	21	120	134	33	41	41	46
1201	170	204	33	200	201	49	49	57	53
2060	235	267	49	291	290	60	60	64	63
3259	292	314	65	354	353	61	61	66	66

TABLE 7.1 – Données TPM, représentées sur les différentes courbes de calibration

### E.2 Données simulation, courbe de calibration

$L_{ADN}$ bp	$R_{  ,Simu}$ $\boxed{2D}$ nm	$R_{ADN}$ $\boxed{3D}$ nm	$R_{ADN,S}$ $\boxed{3D}$ ms	$L_p$ (FJC) nm	$L_p$ (FJC,S) nm	$L_p$ (WLC) nm	$L_p$ (WLC,S) nm
583	154	114	131	33	43	36	49
639	161	128	140	38	45	42	51
1201	216	218	219	58	59	63	64
2060	269	293	293	61	61	64	64
3259	337	385	384	67	67	69	69

TABLE 7.2 – Données simulation, représentées sur les différentes courbes de calibration

$V_{exc}$	$L_{ADN} = 583$				$L_{ADN} = 2060$			
	$V_{exc}$ bp	$V_{exc}$ nm	$L_{p,Simu} = 147bp$	$L_{p,Simu} = 183bp$	$V_{exc}$ bp	$V_{exc}$ nm	$L_{p,Simu} = 147bp$	$L_{p,Simu} = 183bp$
			$R_{  ,Simu}$ $\boxed{2D}$ nm	$R_{  ,Simu}$ $\boxed{2D}$ nm			$R_{  ,Simu}$ $\boxed{2D}$ nm	$R_{  ,Simu}$ $\boxed{2D}$ nm
$a_{mon}$	12	1	154	160	21	7	269	285
$0.6 \cdot a_{mon}$	9	3	154	161	15	3	274	284
$0.4 \cdot a_{mon}$	6	2	155	161	10	4	270	284
$0.2 \cdot a_{mon}$	4	1	154	160	8	3	267	283
$0.1 \cdot a_{mon}$	3	1	154	162	5	2	265	282
$0.05 \cdot a_{mon}$	1	0.5	154	162	3	1	258	280
0	0	0	154	161	0	0	247	267

TABLE 7.3 – Données simulation MC-D, variation du volume exclu

$R_{  ,Simu}$ $\boxed{2D}$ nm	$L_{ADN}$	$L_{p,Simu}$								
		138bp	147bp	157bp	163bp	168bp	173bp	178bp	183bp	188bp
	583 bp	152	154	156	157	158	160	161	162	162
	2060 bp	265	269	274	278	280	282	285	286	286

TABLE 7.4 – Données simulation MC-D, variation de la longueur de persistance

## F DONNÉES SUR L'EFFET DE LA FORCE IONIQUE

$L_p$	$R_{  ,EE}^{2D}$								
	$L_{ADN} = 575$			$L_{ADN} = 583$			$L_{ADN} = 639$		
	[000] nm	[101] nm	[111] nm	[000] nm	[101] nm	[111] nm	[000] nm	[101] nm	[111] nm
107	150.4	142.1	142.2	150.8	143.0	143.1	153.7	149.0	149.2
117	152.2	144.3	144.4	152.7	145.2	145.3	155.8	151.5	151.6
127	153.9	146.4	146.4	154.4	147.5	147.6	157.7	154.2	154.3
137	155.5	148.4	148.4	156.0	149.4	149.4	159.6	156.3	156.4
147	157.0	150.2	150.2	157.6	151.2	151.2	161.3	158.4	158.4
157	158.4	151.7	151.7	159.0	152.7	152.7	163.0	160.1	160.1
167	159.7	153.2	153.2	160.3	154.3	154.3	164.5	161.5	161.5
177	161.0	154.4	154.4	161.1	155.6	155.6	166.0	163.1	163.1
187	162.2	155.6	155.6	162.8	156.8	156.8	167.4	164.4	164.4
197	163.3	156.7	156.7	164.0	158.0	158.0	168.7	165.9	165.9
207	164.3	157.8	157.8	165.1	159.0	159.0	170.0	167.0	167.0

TABLE 7.5 – Données simulation EE, variation du volume exclu

$L_p$	$R_{  ,EE}^{2D}$								
	$L_{ADN} = 1201$			$L_{ADN} = 2060$			$L_{ADN} = 3259$		
	[000] nm	[101] nm	[111] nm	[000] nm	[101] nm	[111] nm	[000] nm	[101] nm	[111] nm
107	179.8	195.4	196.5	213.9	243.2	246.2	253.8	292.8	299.0
117	184.0	200.1	200.9	220.4	250.3	252.8	262.9	301.5	306.9
127	187.9	204.2	204.9	226.7	257.0	259.2	271.6	310.4	315.2
137	191.7	208.2	208.8	232.7	262.8	264.8	280.0	318.8	323.0
147	195.4	212.1	212.6	238.5	269.1	270.8	288.0	326.2	333.0
157	198.9	215.7	216.0	244.1	274.8	276.2	295.8	334.4	337.9
167	202.3	218.9	219.2	249.5	279.7	281.0	303.3	341.2	344.4
177	205.5	222.2	222.5	254.7	284.6	285.7	310.6	348.1	350.9
187	208.6	224.9	225.1	259.8	290.0	291.0	317.7	354.7	357.3
197	211.6	227.5	227.6	264.7	294.6	295.5	324.6	361.5	363.8
207	214.5	230.3	230.4	269.6	298.7	299.5	331.4	367.3	369.4

TABLE 7.6 – Données simulation EE, variation du volume exclu

$C_{te}$	$Fit \ R_{  ,EE}^{2D} = A \times L_p^2 + B \times L_p + C$					
	$L_{ADN} = 575$ nm	$L_{ADN} = 583$ nm	$L_{ADN} = 639$ nm	$L_{ADN} = 1201$ nm	$L_{ADN} = 2060$ nm	$L_{ADN} = 3259$ nm
A	$-6.03 \times 10^{-3}$	$-6.00 \times 10^{-3}$	$-6.83 \times 10^{-3}$	$-9.02 \times 10^{-3}$	$-10.70 \times 10^{-3}$	$-11.52 \times 10^{-3}$
B	1.10	1.11	1.25	1.96	2.71	3.30
C	110.3	110.9	113.0.4	137.3	162.1	193.9

TABLE 7.7 – Données simulation EE, variation du volume exclu

Ion	C <sub>ion</sub> [mmol/L]	I [mmol/L]	ΔI [mmol/L]	1/I [L/mol]	Nb. traj.	τ <sub>  </sub> [ms]	Δτ <sub>  </sub> [ms]	R <sub>exp  </sub> [nm]	ΔR <sub>exp  </sub> [nm]	R <sub>DNA</sub> [nm]	ΔR <sub>DNA</sub> [nm]	L <sub>p</sub> [nm]	ΔL <sub>p</sub> [nm]
Na <sup>+</sup>	0	10.1	0.3	99.4	213	21.2	0.4	150.5	0.5	127.0	1.9	91.0	5.6
	5	15.1	0.3	66.4	125	21.3	0.5	160.7	0.7	127.4	2.7	91.0	6.2
	10	20.1	0.4	49.9	144	21.2	0.4	159.0	0.7	124.2	2.6	76.7	10.5
	15	25.1	0.5	39.9	84	21.2	0.3	159.8	0.8	125.7	1.5	82.4	4.6
	20	30.1	0.5	33.3	117	20.8	0.3	159.8	0.7	125.7	2.7	82.4	8.7
	25	35.1	0.7	28.5	86	20.4	0.3	159.6	0.9	125.3	3.4	80.7	10.0
	30	40.1	0.7	25.0	90	20.8	0.3	159.6	0.8	125.3	3.0	80.7	9.7
	40	50.1	0.8	20.0	106	20.8	0.4	158.8	0.8	123.8	3.2	75.6	11.3
	50	60.1	1.2	16.7	84	20.1	01.1	159	0.8	124.2	3.0	75.7	10.8
	100	110.1	1.5	9.1	94	19.7	0.4	156.7	0.8	119.7	3.2	66.8	6.2
	155	165.1	2.7	6.1	101	19.5	0.3	154.8	0.8	116.0	3.2	61.0	4.5
	500	510.1	7.7	2.0	-	-	-	-	-	-	-	-	-
	1500	1510.1	20.7	0.7	-	-	-	-	-	-	-	-	-
	3000	3010.1	37.5	0.3	-	-	-	-	-	-	-	-	-
Mg <sup>2+</sup>	0.15	10.5	0.3	95.1	-	-	-	-	-	-	-	-	-
	0.25	10.8	0.3	92.5	-	-	-	-	-	-	-	-	-
	0.50	11.6	0.3	86.5	97	19.7	0.4	156.1	0.9	118.5	3.6	64.8	5.9
	0.75	12.3	0.3	81.2	-	-	-	-	-	-	-	-	-
	1	13.1	0.3	76.6	106	20.6	0.3	152.4	0.7	111.1	2.9	55.0	3.2
	2	16.1	0.4	62.3	90	19.6	0.3	152.9	0.9	116.2	3.6	61.3	5.1
	5	25.1	0.6	39.9	56	18.3	0.3	151.7	1.0	109.6	4.2	53.4	4.5
	10	40.1	0.7	25.0	-	-	-	-	-	-	-	-	-
	20	70.1	1.2	14.3	80	18.8	0.3	150.4	0.9	106.9	3.8	50.6	3.6
	50	160.1	2.7	6.2	66	18.9	0.4	151.4	1.0	109.0	4.2	52.7	4.4
	100	310.1	5.1	3.2	69	19.6	0.6	149.9	0.9	104.8	4.8	48.6	3.6

TABLE 7.8 – **ADN de 575 bp**. Conditions saline (concentration en ions ajoutée, force ionique et barre d'erreur associée), Nombre de trajectoires TPM, temps de corrélation expérimental  $\tau_{||}$  et barre d'erreur associée et résultats expérimentaux de la mesure de la distance bout-à-bout des complexes ADN/particules,  $R_{\text{exp||}}$ , avec barre d'erreur associée. Distance bout-à-bout de la molécule d'ADN seule correspondantes,  $R_{\text{DNA}}$  avec barre d'erreur associée. Longueur de persistance extraite via la résolution du problème inverse,  $L_p$ , avec barre d'erreur associée.

Ion	$C_{\text{ion}}$ [mmol/L]	$I$ [mmol/L]	$\Delta I$ [mmol/L]	$1/I$ [L/mol]	Nb. traj.	$\tau_{  }$ [ms]	$\Delta\tau_{  }$ [ms]	$R_{\text{exp}  }$ [nm]	$\Delta R_{\text{exp}  }$ [nm]	$R_{\text{DNA}}$ [nm]	$\Delta R_{\text{DNA}}$ [nm]	$L_p$ [nm]	$\Delta L_p$ [nm]
$\text{Na}^+$	0	10.1	0.3	99.4	1115	19.5	0.2	152.8	0.3	111.9	1.2	53.4	1.3
	5	15.1	0.3	66.4	409	19.4	0.3	152.4	0.5	111.1	2.0	52.6	2.1
	10	20.1	0.4	49.9	664	19.3	0.2	152.9	0.4	112.1	1.6	53.6	1.8
	15	25.1	0.5	39.9	438	19.2	0.3	152.0	0.5	110.3	2.0	51.7	2.1
	20	30.1	0.5	33.3	707	20.4	0.2	153.7	0.3	113.7	1.2	55.4	1.3
	25	35.1	0.7	28.5	698	19.7	0.2	151.8	0.4	109.8	1.2	51.3	1.2
	30	40.1	0.7	25.0	438	20.0	0.2	153.4	0.5	113.1	1.6	54.7	1.6
	40	50.1	0.8	20.0	365	18.7	0.2	154.5	0.4	115.3	1.0	57.3	2.2
	50	60.1	1.2	16.7	820	19.5	0.2	153.1	0.4	112.5	1.7	54.1	1.9
	100	110.1	1.5	9.1	626	19.2	0.2	152.1	0.4	110.5	1.7	51.9	1.7
	155	165.1	2.7	6.1	746	18.8	0.2	151.2	0.4	108.6	1.7	50.1	1.6
	500	510.1	7.7	2.0	210	19.3	0.2	148.8	0.6	103.5	2.6	45.6	2.2
	1500	1510.1	20.7	0.7	180	18.9	0.2	147.6	0.6	100.9	2.6	43.5	2.0
	3000	3010.1	37.5	0.3	-	-	-	-	-	-	-	-	-
$\text{Mg}^{2+}$	0.15	10.5	0.3	95.1	218	19.4	0.3	150.6	0.6	107.3	2.1	48.9	2.3
	0.25	10.8	0.3	92.5	218	18.1	0.3	147.1	0.6	99.8	2.9	42.6	2.1
	0.50	11.6	0.3	86.5	375	17.9	0.3	147.6	0.5	100.9	4.0	43.5	1.7
	0.75	12.3	0.3	81.2	65	17.6	0.4	149.2	1.2	104.4	1.0	46.3	4.4
	1	13.1	0.3	76.6	947	18.0	0.2	147.9	0.3	101.5	1.6	44.0	1.0
	2	16.1	0.4	62.3	395	17.2	0.3	146.4	0.5	98.2	1.5	41.5	1.7
	5	25.1	0.6	39.9	322	17.0	0.3	145.7	0.5	96.7	2.6	40.3	1.6
	10	40.1	0.7	25.0	782	16.6	0.2	144.2	0.9	93.0	4.9	37.9	2.8
	20	70.1	1.2	14.3	279	16.4	0.4	142.9	1.0	90.2	2.7	36.0	3.0
	50	160.1	2.7	6.2	270	16.9	0.4	143.2	0.6	90.9	1.8	36.4	1.7
	100	310.1	5.1	3.2	142	20.8	0.6	141.0	0.7	85.6	1.8	33.2	2.0

TABLE 7.9 – **ADN de 583 bp**. Conditions saline (concentration en ions ajoutée, force ionique et barre d'erreur associée), Nombre de trajectoires TPM, temps de corrélation expérimental  $\tau_{||}$  et barre d'erreur associée et résultats expérimentaux de la mesure de la distance bout-à-bout des complexes ADN/particules  $R_{\text{exp}||}$ , avec barre d'erreur associée. Distance bout-à-bout de la molécule d'ADN seule correspondantes,  $R_{\text{DNA}}$  avec barre d'erreur associée. Longueur de persistance extraite via la résolution du problème inverse,  $L_p$ , avec barre d'erreur associée.

Ion	C <sub>ion</sub> [mmol/L]	I [mmol/L]	ΔI [mmol/L]	1/I [L/mol]	Nb. traj.	τ <sub>  </sub> [ms]	Δτ <sub>  </sub> [ms]	R <sub>exp  </sub> [nm]	ΔR <sub>exp  </sub> [nm]	R <sub>DNA</sub> [nm]	ΔR <sub>DNA</sub> [nm]	L <sub>p</sub> [nm]	ΔL <sub>p</sub> [nm]
Na <sup>+</sup>	0	10.1	0.3	99.4	496	19.1	0.2	149.5	0.4	105.0	1.7		
	5	15.1	0.3	66.4	199	18.0	0.5	150.8	0.7	107.8	1.7	64.8	1.5
	10	20.1	0.4	49.9	294	19.0	0.4	150.8	0.5	107.8	2.9	49.3	2.7
	15	25.1	0.5	39.9	153	17.6	0.5	152.9	0.8	112.1	2.0	53.6	2.2
	20	30.1	0.5	33.3	-	-	-	-	-	-	-	-	-
	25	35.1	0.7	28.5	283	19.3	0.5	151.8	0.5	109.8	3.3	51.3	3.3
	30	40.1	0.7	25.0	149	18.6	0.7	150.9	0.8	108.0	2.1	49.5	2.0
	40	50.1	0.8	20.0	152	18.9	0.8	150.6	0.9	107.3	3.4	48.9	3.1
	50	60.1	1.2	16.7	251	18.6	0.6	149.0	0.6	103.9	3.9	45.9	3.3
	100	110.1	1.5	9.1	288	18.8	0.4	148.0	0.5	101.8	2.6	44.2	2.1
	155	165.1	2.7	6.1	287	17.5	0.5	145.0	0.6	96.4	2.3	40.2	1.6
	500	510.1	7.7	2.0	-	-	-	-	-	-	-	-	-
	1500	1510.1	20.7	0.7	-	-	-	-	-	-	-	-	-
	3000	3010.1	37.5	0.3	-	-	-	-	-	-	-	-	-
	0.15	10.5	0.3	95.1	-	-	-	-	-	-	-	-	-
	0.25	10.8	0.3	92.5	-	-	-	-	-	-	-	-	-
Mg <sup>2+</sup>	0.50	11.6	0.3	86.5	123	16.9	0.7	146.0	0.9	97.3	4.1	40.8	2.9
	0.75	12.3	0.3	81.2	-	-	-	-	-	-	-	-	-
	1	13.1	0.3	76.6	187	17.2	0.5	146.7	0.7	98.9	3.1	41.9	2.3
	2	16.1	0.4	62.3	91	17.0	1.1	144.1	0.8	93.0	5.6	37.8	3.7
	5	25.1	0.6	39.9	122	15.0	0.9	142.0	0.8	88.0	3.9	34.7	2.3
	10	40.1	0.7	25.0	105	16.1	0.4	141.5	0.8	86.0	3.9	33.9	2.3
	20	70.1	1.2	14.3	111	16.0	0.5	142.5	0.8	89.2	3.8	35.4	2.4
	50	160.1	2.7	6.2	-	-	-	-	-	-	-	-	-
	100	310.1	5.1	3.2	110	16.0	0.4	140.9	0.8	85.3	4.0	33.1	2.3

TABLE 7.10 – ADN de 583 bp avec une PolyA de 50 bp. Conditions saline (concentration en ions ajoutée, force ionique et barre d’erreur associée), Nombre de trajectoires TPM, temps de corrélation expérimental  $\tau_{||}$  et barre d’erreur associée et résultats expérimentaux de la mesure de la distance bout-à-bout des complexes ADN/particules,  $R_{\text{exp}}||$ , avec barre d’erreur associée. Distance bout-à-bout de la molécule d’ADN seule correspondantes,  $R_{\text{DNA}}$  avec barre d’erreur associée. Longueur de persistance extraite via la résolution du problème inverse,  $L_p$ , avec barre d’erreur associée.



Ion	$C_{\text{ion}}$ [mmol/L]	$I$ [mmol/L]	$\Delta I$ [mmol/L]	$1/I$ [L/mol]	Nb. traj.	$\tau_{\parallel}$ [ms]	$\Delta\tau_{\parallel}$ [ms]	$R_{\text{exp}\parallel}$ [nm]	$\Delta R_{\text{exp}\parallel}$ [nm]	$R_{\text{DNA}}$ [nm]	$\Delta R_{\text{DNA}}$ [nm]	$L_p$ [nm]	$\Delta L_p$ [nm]
$\text{Na}^+$	0	10.1	0.3	99.4	292	20.70	0.2	158.4	0.5	123.0	1.6	50.3	1.8
	5	15.1	0.3	66.4	159	20.6	0.4	157.5	0.6	121.3	2.3	48.7	2.1
	10	20.1	0.4	49.9	172	20.4	0.3	159.1	0.6	124.4	2.3	51.6	2.2
	15	25.1	0.5	39.9	165	20.7	0.3	155.4	0.8	117.1	3.2	45.2	2.6
	20	30.1	0.5	33.3	160	21.0	0.4	157.0	0.6	120.3	2.3	47.9	2.0
	25	35.1	0.7	28.5	123	21.1	0.3	156.9	0.6	120.1	2.4	47.7	2.0
	30	40.1	0.7	25.0	139	19.9	0.3	154.9	0.7	116.2	2.8	44.4	2.2
	40	50.1	0.8	20.0	170	20.3	0.6	155.7	0.9	117.7	3.6	45.7	2.9
	50	60.1	1.2	16.7	154	20.2	0.3	154.4	0.7	115.1	2.8	43.7	2.2
	100	110.1	1.5	9.1	178	20.1	0.4	153.6	0.6	113.5	2.4	42.5	1.8
	155	165.1	2.7	6.1	122	19.5	0.4	155.5	0.9	117.3	3.6	45.4	2.9
	500	510.1	7.7	2.0	-	-	-	-	-	-	-	-	-
	1500	1510.1	20.7	0.7	-	-	-	-	-	-	-	-	-
	3010.1	3010.1	37.5	0.3	-	-	-	-	-	-	-	-	-
$\text{Mg}^{2+}$	0.15	10.5	0.3	95.1	-	-	-	-	-	-	-	-	-
	0.25	10.8	0.3	92.5	-	-	-	-	-	-	-	-	-
	0.50	11.6	0.3	86.5	1205	19.4	0.3	153.0	0.7	112.3	2.9	41.6	2.1
	0.75	12.3	0.3	81.2	-	-	-	-	-	-	-	-	-
	1	13.1	0.3	76.6	112	18.2	0.3	152.2	0.8	110.7	3.3	40.4	2.3
	2	16.1	0.4	62.3	94	18.5	0.4	151.5	0.9	109.2	3.7	39.4	2.5
	5	25.1	0.6	39.9	88	18.6	0.7	153.2	1.0	112.7	4.1	41.9	3.0
	10	40.1	0.7	25.0	111.0	18.0	0.7	149.3	0.8	104.6	3.4	36.4	2.1
	20	70.1	1.2	14.3	105	18.2	0.2	150.9	0.7	108.0	2.9	38.6	1.9
	50	160.1	2.7	6.2	130	18.2	0.3	148.4	0.7	104.8	3.0	36.5	1.9
	100	310.1	5.1	3.2	105	18.8	0.5	148.5	0.9	102.9	3.9	35.3	2.4

TABLE 7.11 – **ADN de 639 bp**. Conditions saline (concentration en ions ajoutée, force ionique et barre d'erreur associée), Nombre de trajectoires TPM, temps de corrélation expérimental  $\tau_{\parallel}$  et barre d'erreur associée et résultats expérimentaux de la mesure de la distance bout-à-bout des complexes ADN/particules  $R_{\text{exp}\parallel}$ , avec barre d'erreur associée. Distance bout-à-bout de la molécule d'ADN seule correspondantes,  $R_{\text{DNA}}$  avec barre d'erreur associée. Longueur de persistance extraite via la résolution du problème inverse,  $L_p$ , avec barre d'erreur associée.

Ion	C <sub>ion</sub> [mmol/L]	I [mmol/L]	ΔI [mmol/L]	1/I [L/mol]	Nb. traj.	τ <sub>  </sub> [ms]	Δτ <sub>  </sub> [ms]	R <sub>exp  </sub> [nm]	ΔR <sub>exp  </sub> [nm]	R <sub>DNA</sub> [nm]	ΔR <sub>DNA</sub> [nm]	L <sub>p</sub> [nm]	ΔL <sub>p</sub> [nm]
Na <sup>+</sup>	0	10.1	0.3	99.4	1399	35.8	0.2	213.6	0.2	216.6	0.3	51.0	0.1
	5	15.1	0.3	66.4	215	35.9	0.7	214.9	0.5	216.3	1.5	52.3	1.0
	10	20.1	0.4	49.9	250	35.1	0.5	213.6	0.6	214.3	1.8	51.0	1.2
	15	25.1	0.5	39.9	232	35.4	0.4	213.0	0.5	213.4	1.5	50.4	1.0
	20	30.1	0.5	33.3	231	34.5	0.5	213.4	0.5	214.0	1.5	50.8	1.0
	25	35.1	0.7	28.5	269	34.3	0.5	212.7	0.5	213.0	1.5	50.1	1.0
	30	40.1	0.7	25.0	216	35.2	0.7	211.5	0.6	211.2	1.8	49.0	1.1
	40	50.1	0.8	20.0	55	35.4	0.7	211.7	1.0	211.5	3.0	49.2	1.9
	50	60.1	1.2	16.7	270	33.5	0.2	209.6	0.3	208.3	0.9	47.3	0.6
	100	110.1	1.5	9.1	273	33.0	0.3	208.4	0.5	206.5	1.5	46.2	0.9
	155	165.1	2.7	6.1	869	32.7	0.2	200.5	0.3	194.4	1.0	39.5	0.5
	500	510.1	7.7	2.0	587	32.5	0.2	196.0	0.3	187.4	1.0	36.0	0.5
	1500	1510.1	20.7	0.7	569	35.4	0.2	192.1	0.3	181.3	1.0	33.0	0.5
	3000	3010.1	37.5	0.3	587	39.9	0.2	189.9	0.3	177.6	01.0	31.4	0.5
Mg <sup>2+</sup>	0.15	10.5	0.3	95.1	42	33.2	1.3	208.4	1.5	206.5	4.6	46.2	2.7
	0.25	10.8	0.3	92.0	33.6	34	0.8	204.8	1.7	201.0	5.2	43.1	2.9
	0.5	11.6	0.3	86.5	714	33.5	0.2	206.4	0.2	203.5	0.6	44.4	0.4
	0.75	12.3	0.3	81.2	83	32.1	0.4	204.7	0.9	200.9	2.8	43.0	1.6
	1	13.1	0.3	76.6	685	32.8	0.2	205.8	0.3	202.6	0.9	43.9	0.5
	2	16.1	0.4	62.3	766	32.1	0.2	203.4	0.3	198.9	0.9	41.9	0.5
	5	25.1	0.6	39.9	636	31.8	0.2	202.3	0.3	197.2	0.9	41.0	0.5
	10	40.1	0.7	25.0	499	31.1	0.3	200.9	0.3	195.0	1.0	39.8	0.5
	20	70.1	1.2	14.3	443	30.7	0.3	209.8	0.3	193.8	1.0	38.9	0.5
	50	160.1	2.7	6.2	-	-	-	-	-	-	-	-	-
	100	310.1	5.1	3.2	348	31.5	0.3	197.0	0.4	189.1	1.3	36.7	0.6

TABLE 7.12 – **ADN de 1201 bp.** Conditions saline (concentration en ions ajoutée, force ionique et barre d'erreur associée), Nombre de trajectoires TPM, temps de corrélation expérimental  $\tau_{||}$  et barre d'erreur associée et résultats expérimentaux de la mesure de la distance bout-à-bout des complexes ADN/particules,  $R_{exp||}$ , avec barre d'erreur associée. Distance bout-à-bout de la molécule d'ADN seule correspondantes,  $R_{DNA}$  avec barre d'erreur associée. Longueur de persistance extraite via la résolution du problème inverse,  $L_p$ , avec barre d'erreur associée.

Ion	$C_{\text{ion}}$ [mmol/L]	$I$ [mmol/L]	$\Delta I$ [mmol/L]	$1/I$ [L/mol]	Nb. traj.	$\tau_{  }$ [ms]	$\Delta\tau_{  }$ [ms]	$R_{\text{exp}  }$ [nm]	$\Delta R_{\text{exp}  }$ [nm]	$R_{\text{DNA}}$ [nm]	$\Delta R_{\text{DNA}}$ [nm]	$L_p$ [nm]	$\Delta L_p$ [nm]
$\text{Na}^+$	0	10.1	0.3	99.4	1115	51.6	0.3	276.5	0.2	303.6	0.2	53.7	0.2
	5	15.1	0.3	66.4	204	52.8	0.6	276.0	0.7	302.9	1.9	53.4	0.9
	10	20.1	0.4	49.9	284	52.0	0.5	278.2	0.5	305.9	1.4	54.8	0.7
	15	25.1	0.5	39.9	197	51.1	0.6	277.0	0.6	304.3	1.7	54.0	0.8
	20	30.1	0.5	33.3	230	52.2	0.6	275.6	0.5	302.4	1.7	53.1	0.7
	25	35.1	0.7	28.5	312	51.2	0.6	276.0	0.5	303.1	1.4	53.4	0.7
	30	40.1	0.7	25.0	193	52.0	0.7	272.0	0.7	297.4	1.9	50.8	0.9
	40	50.1	0.8	20.0	223	51.8	0.5	271.4	0.6	296.6	1.7	50.5	0.8
	50	60.1	1.2	16.7	299	51.2	0.2	270.6	0.5	295.5	1.4	50.0	0.6
	100	110.1	1.5	9.1	254	50.2	0.4	267.0	0.6	290.6	1.7	47.8	0.7
	155	165.1	2.7	6.1	1100	49.9	0.2	254.1	0.3	272.7	0.9	40.5	0.4
	500	510.1	7.7	2.0	732	48.6	0.3	251.0	0.3	268.3	0.9	38.8	0.3
	1500	1510.1	20.7	0.7	840	50.9	0.2	248.1	0.3	264.3	0.9	37.3	0.3
	3000	3010.1	37.5	0.3	815	57.1	0.3	243.6	0.2	257.9	0.6	35.0	0.5
$\text{Mg}^{2+}$	0.15	10.5	0.3	95.1	364	49.8	0.3	270.9	0.4	295.9	1.1	50.2	0.5
	0.25	10.8	0.3	92.5	197	48.9	0.4	265.2	0.5	288.1	1.4	46.8	0.6
	0.50	11.6	0.3	86.5	482	48.0	0.4	267.7	0.3	291.5	0.9	48.2	0.4
	0.75	12.3	0.3	81.2	349	47.4	0.3	267.7	0.4	291.5	1.1	48.2	0.5
	1	13.1	0.3	76.6	559	47.9	0.3	266.3	0.3	189.6	0.9	47.4	0.6
	2	16.1	0.4	62.3	524	47.2	0.4	263.7	0.3	286.0	0.9	45.9	0.4
	5	25.1	0.6	39.9	188	46.0	0.7	260.6	0.5	181.7	1.4	44.1	0.4
	10	40.1	0.7	25.0	184	46.4	0.7	255.8	0.9	275.0	2.5	41.4	1.0
	20	70.1	1.2	14.3	156	46.0	0.6	254.3	0.8	273.0	2.3	40.6	0.9
	50	160.1	2.7	6.2	-	-	-	-	-	-	-	-	-
	100	310.1	5.1	3.2	118	45.5	0.6	248.0	1.2	264.1	3.4	37.2	1.3

TABLE 7.13 – **ADN de 2060 bp**. Conditions saline (concentration en ions ajoutée, force ionique et barre d'erreur associée), Nombre de trajectoires TPM, temps de corrélation expérimental  $\tau_{||}$  et barre d'erreur associée et résultats expérimentaux de la mesure de la distance bout-à-bout des complexes ADN/particules  $R_{\text{exp}||}$ , avec barre d'erreur associée. Distance bout-à-bout de la molécule d'ADN seule correspondantes,  $R_{\text{DNA}}$  avec barre d'erreur associée. Longueur de persistance extraite via la résolution du problème inverse,  $L_p$ , avec barre d'erreur associée.

Ion	C <sub>ion</sub> [mmol/L]	I [mmol/L]	ΔI [mmol/L]	1/I [L/mol]	Nb. traj.	τ <sub>  </sub> [ms]	Δτ <sub>  </sub> [ms]	R <sub>exp  </sub> [nm]	ΔR <sub>exp  </sub> [nm]	R <sub>DNA</sub> [nm]	ΔR <sub>DNA</sub> [nm]	L <sub>p</sub> [nm]	ΔL <sub>p</sub> [nm]
Na <sup>+</sup>	0	10.1	0.3	99.4	195	72.1	0.7	337.2	0.8	376.5	2.1	53.2	0.8
	5	15.1	0.3	66.4	107	70.6	0.9	330.8	0.9	384.8	2.4	50.3	0.8
	10	20.1	0.4	49.9	65	70.5	1.1	328.3	1.7	373.1	4.5	49.1	1.6
	15	25.1	0.5	39.9	50	70.5	1.4	332.2	1.4	378.2	3.7	50.9	1.3
	20	30.1	0.5	33.3	80	70.2	1.1	332.4	1.1	378.5	2.9	51.0	1.0
	25	35.1	0.7	28.5	81	68.5	1.1	329.5	1.1	374.6	2.9	49.6	1.0
	30	40.1	0.7	25.0	80	68.8	1.1	331.2	1.1	376.9	2.9	50.4	1.0
	40	50.1	0.8	20.0	85	68.4	0.9	327.9	1.1	372.5	2.9	48.9	1.0
	50	60.1	1.2	16.7	103	68.6	0.9	324.6	1.2	368.2	3.2	47.4	1.1
	100	110.1	1.5	9.1	71	67.8	1.3	323.3	1.2	366.4	3.2	46.8	1.1
	155	165.1	2.7	6.1	90	65.6	1.1	322.2	0.9	365.0	2.4	46.3	0.8
	500	510.1	7.7	2.0	-	-	-	-	-	-	-	-	-
	1500	1510.1	20.7	0.7	-	-	-	-	-	-	-	-	-
	3000	3010.1	37.5	0.3	-	-	-	-	-	-	-	-	-
Mg <sup>2+</sup>	0.15	10.5	0.3	95.1	-	-	-	-	-	-	-	-	-
	0.25	10.8	0.3	92.5	-	-	-	-	-	-	-	-	-
	0.50	11.6	0.3	86.5	15	61.7	2.6	319.8	2.1	361.8	5.6	45.2	1.9
	0.75	12.3	0.3	81.2	-	-	-	-	-	-	-	-	-
	1	13.1	0.3	76.6	107	64.2	0.8	316.5	0.8	357.4	2.1	43.8	0.7
	2	16.1	0.4	62.3	61	63.0	1.1	313.4	1.0	353.3	2.7	42.4	0.9
	5	25.1	0.6	39.9	59	61.2	1.4	309.1	1.1	347.6	2.9	40.6	0.9
	10	40.1	0.7	25.0	70	59.6	0.9	302.6	1.0	338.9	2.7	37.9	0.8
	20	70.1	1.2	14.3	73	59.1	0.8	302.3	1.1	338.5	2.9	37.8	0.9
	50	160.1	2.7	6.2	45	59.5	1.2	303.2	1.2	339.7	3.2	38.1	1.90
	100	310.1	5.1	3.2	62	59.2	0.9	302.4	0.9	338.6	2.4	37.8	0.7

TABLE 7.14 – **ADN de 3259 bp**. Conditions saline (concentration en ions ajoutée, force ionique et barre d'erreur associée), Nombre de trajectoires TPM, temps de corrélation expérimental  $\tau_{||}$  et barre d'erreur associée et résultats expérimentaux de la mesure de la distance bout-à-bout des complexes ADN/particules,  $R_{\text{exp||}}$ , avec barre d'erreur associée. Distance bout-à-bout de la molécule d'ADN seule correspondantes,  $R_{\text{DNA}}$  avec barre d'erreur associée. Longueur de persistance extraite via la résolution du problème inverse,  $L_p$ , avec barre d'erreur associée.

## G DONNÉES SUR L'EFFET D'UNE COURBURE LOCALE

Se référer aux tableaux de données contenue dans l'article.

## H DONNÉES SUR L'EFFET DE LA TEMPÉRATURE

$L_{ADN} = 583 \text{ bp}$								
$T$ °C	$\Delta T$ °C	Nbr Traj.	$\tau_{  }$ ms	$\Delta\tau_{  }$ ms	$R_{Exp  }$ nm	$\Delta R_{Exp  }$ nm	$L_p$ nm	$\Delta L_p$ nm
15	0.9	491	23.3	0.2	111.7	0.4	-	-
20	0.5	429	20.2	0.2	107.6	0.4	-	-
25	0.1	462	17.2	0.1	103.4	0.4	-	-
30	< 0.01	478	15.6	0.2	100.3	0.5	-	-
35	0.3	506	14.3	0.2	97.9	0.5	-	-
40	0.6	443	-	-	-	-	-	-
45	1.1	493	-	-	-	-	-	-
50	1.3	378	-	-	-	-	-	-
55	1.9	338	-	-	-	-	-	-
60	2.2	206	-	-	-	-	-	-
65	2.5	247	-	-	-	-	-	-
70	3.8	143	-	-	-	-	-	-
75	3.6	78	-	-	-	-	-	-
80	4.4	25	-	-	-	-	-	-

$L_{ADN, Poly(dA)} = 583 \text{ bp}$								
$T$ °C	$\Delta T$ °C	Nbr Traj.	$\tau_{  }$ ms	$\Delta\tau_{  }$ ms	$R_{Exp  }$ nm	$\Delta R_{Exp  }$ nm	$L_p$ nm	$\Delta L_p$ nm
15	0.9	452	22.6	0.2	110.0	0.5	-	-
20	0.5	365	19.4	0.2	107.6	0.4	-	-
25	0.1	310	16.9	0.2	103.9	0.5	-	-
30	< 0.01	203	14.9	0.3	100.7	0.7	-	-
35	0.3	156	-	-	-	-	-	-
40	0.6	216	-	-	-	-	-	-
45	1.1	146	-	-	-	-	-	-
50	1.3	111	-	-	-	-	-	-
55	1.9	81	-	-	-	-	-	-
60	2.2	32	-	-	-	-	-	-
65	2.5	18	-	-	-	-	-	-
70	3.8	4	-	-	-	-	-	-
75	3.6	-	-	-	-	-	-	-
80	4.4	-	-	-	-	-	-	-

## Titre de la thèse :

Annaëlle  
BRUNET  
1, chemin Reboul  
31100  
TOULOUSE  
(FRANCE)

**Étude à l'échelle de la molécule unique des changements  
conformationnels de la molécule d'ADN.**  
***Influence de la présence de défauts locaux présents sur l'ADN et  
de paramètres physico-chimiques de la solution environnante***

+33 (0)6 87 67 58 42  
[annael.brunet@live.fr](mailto:annael.brunet@live.fr)  
WebSite :  
Abrunet.com

## Résumé français de la thèse :

Les ions jouent un rôle majeur sur les processus biologiques affectant la molécule d'ADN que ce soit en termes d'activité de liaison de protéines à l'ADN ou d'encapsulation de l'ADN dans les capsides virales ou le noyau. L'activité de protéines sur l'ADN est, par ailleurs, fréquemment liée à une courbure locale de l'axe de la double hélice, que ce soit en raison d'une séquence intrinsèquement courbée, ou, via la capacité de protéines à courber la séquence sur laquelle elles se fixent. Être capable de caractériser et comprendre l'effet des ions présents en solution, de la courbure et de la dénaturation locale de la molécule d'ADN sur les conformations de cette dernière est donc crucial pour approfondir la compréhension de nombreux processus biologiques. Des travaux, tant expérimentaux que théoriques, ont déjà été menés sur ces questions mais celles-ci sont encore largement débattues. En effet, pour y répondre, doivent notamment être développées des méthodes expérimentales qui ne perturbent pas significativement la conformation de l'ADN ou le complexe ADN-protéine, ainsi que des modèles théoriques associés permettant une analyse précise des données expérimentales et leur compréhension physique.

L'objectif de ce travail est de proposer des outils expérimentaux et théoriques permettant de décrire physiquement l'influence de défauts locaux présents sur la molécule d'ADN et de paramètres physico-chimiques de la solution environnante.

A cette fin, des données expérimentales ont été acquises à l'échelle de la molécule unique grâce à la technique haut-débit de "Tethered Particle Motion" (HT-TPM). Le TPM consiste à enregistrer, au cours du temps, les positions d'une particule accrochée à l'extrémité d'une molécule d'ADN, immobilisée par son autre extrémité sur un support en verre. L'utilisation d'une biopuce permettant la parallélisation des complexes ADN/particule et l'acquisition "à haut débit" de données TPM a permis d'obtenir une grande accumulation de statistiques individuelles. Une procédure d'analyse efficace a été élaborée afin de déterminer les amplitudes du mouvement des assemblages ADN-particules valides. En parallèle, ont été effectuées des simulations basées sur un modèle de physique statistique mésoscopique dans lequel la molécule d'ADN est assimilée à une chaîne de billes de rayons variables dont les déplacements sont régis par la diffusion brownienne et une énergie potentielle d'interaction prenant en compte notamment l'énergie de courbure du polymère ADN.

Une première étude a porté sur l'effet de la force ionique de la solution environnante sur la longueur de persistance  $L_p$ , qui traduit la rigidité du polymère d'ADN. Les valeurs de  $L_p$  extraites des données de HT-TPM ont fait apparaître une décroissance de la longueur de persistance de 55 à 30 nm, corrélée à l'augmentation de la force ionique, avec une décroissance plus forte observée pour les ions divalents  $Mg^{2+}$  que pour les ions monovalents  $Na^+$ . Les valeurs de  $L_p$  déterminées sur une plage étendue de force ionique ont permis de valider l'approche théorique proposée par Manning en 2006 dans la cas  $Na^+$ .

Une deuxième étude a conduit à l'élaboration d'une méthode permettant de quantifier l'angle de courbure locale induite par une séquence spécifique ou la liaison d'une protéine sur la molécule d'ADN. L'échantillon modèle a été obtenu en insérant de une à sept séquences CAAAAACGG en phase. Une description théorique de la chaîne d'ADN appelée "kinked Worm-Like Chain" a été proposée. Elle conduit à une formule simple de la distance bout-à-bout de l'ADN qui permet d'extraire la valeur de l'angle de courbure à partir des mesures de HT-TPM. Ainsi, il a pu être montré que la séquence CAAAAACGG induit un angle de  $19^\circ \pm 4^\circ$  en accord avec les données de la littérature.

Une troisième étude concernant la mesure de l'impact de la dénaturation partielle de l'ADN, induite par la température, sur sa rigidité apparente globale a été menée. Des résultats préliminaires sont proposés.

#### **Résumé anglais de la thèse :**

Ions play an important role in many biological processes affecting the DNA molecule, both for binding activities of DNA-protein interaction, and the DNA packaging in viral capsids or in the cell nucleus. Proteins actions on DNA are also often associated to the double helix curvature, be it because of an intrinsic curved sequence, or of the ability of the proteins, to curve the sequence they are trying to bind. Being able to characterize and understand the effects on the DNA conformation of ions present in solution, DNA local curvature, and local denaturation bubble is essential and crucial for the thorough understanding of many biological processes. Many experimental, and theoretical studies have already been conducted to address these questions. However they remain highly debated. To answer then one must notably develop experimental approaches that minimize alteration of the conformation of the DNA molecule or the complex protein-DNA, as well as associated theoretical models that permit a precise analysis of experimental data as well as their physical understanding.

The goal of this work is to develop and propose experimental and theoretical tools which would provide a physical description of the influence of DNA local defects on the DNA molecule as well as of physicochemical conditions of the DNA environmental solution.

For this purpose, experimental data have been collected, at a single molecule level, using the High-Throughput Tethered Particle Motion" (HT-TPM) technique. TPM consists of recording the location of a particle grafted by one end of a single DNA molecule and immobilized, at the other end, to a glass surface. The use of a biochip that enables the parallelization of DNA/particle complexes and the ensuing high-throughput data acquisition permitted to obtain a large accumulation of individual statistics. A strong analysis procedure has been developed to extract and quantify the amplitude of motion of the valid DNA/particle complexes. Alongside that, simulations have been run, based on a mesoscopic statistical mechanics model in which the DNA molecule is related to a chain of monomers with varying radius and in which the amplitude of motion is governed by both the Brownian motion and by the interaction potential associated to stretching and bending energies of the polymer.

A first study was conducted on the effect of the ionic strength induced by surrounding ions in solution on the DNA persistence length ( $L_p$ ) which characterizes the DNA polymer rigidity. The extracted  $L_p$  values of HT-TPM measurement decrease from 55 to 30 nm when the ionic strength increases. A stronger decrease was observed in presence of divalent ions  $Mg^{2+}$  than with monovalent ions  $Na^+$ . This quantification of  $L_p$  dependence, on a large and strongly prospected range of ionic strengthes, tends to validate the theoretical approach proposed in 2006 by Manning in presence of monovalent ions  $Na^+$ .

A second project allows us to develop a method of evaluation and quantification of local DNA bending angles, induced either by specific intrinsic sequence, or by the binding of proteins on DNA. Constructs made of 575 base-pair DNAs with in-phase assemblies of one to seven sequences CAAAAACGG was used. A theoretical description of the polymer chain, named "kinked Worm-Like Chain" was proposed which leads to a simple formulation of the end-to-end distance of DNA molecules allowing to extract local bend angles from HT-TPM measurement. As a result, we find that the sequence CAAAAACGG induces a bend angle of  $19^\circ \pm 4^\circ$  in agreement with other value from the literature.

A third study concerning the influence of temperature-induced partial denaturation on the global apparent rigidity parameters of the polymer was conducted. Preliminary results are proposed.

#### **Titre anglais de la thèse :**

Study at the single molecule level of conformational changes of the DNA molecule. *Impact of local defects included in the DNA molecule and of a large set of physicochemical conditions*

#### **Mots clés de la thèse :**

Molécule unique, Polymère, ADN, "Tethered Particle Motion", Longueur de persistance, Simulation Monte-Carlo, Physique statistique



$L_{ADN} = 1201 \text{ bp}$								
$T$ °C	$\Delta T$ °C	Nbr Traj.	$\tau_{  }$ ms	$\Delta\tau_{  }$ ms	$R_{Exp  }$ nm	$\Delta R_{Exp  }$ nm	$L_p$ nm	$\Delta L_p$ nm
15	0.9	895	40.1	0.1	172.3	0.2	39.8	0.7
20	0.5	600	35.5	0.2	167.3	0.3	38.1	1.0
25	0.1	680	31.0	0.1	165.1	0.3	39.4	0.9
30	< 0.01	520	27.3	0.1	159.1	0.4	37.4	1.3
35	0.3	669	24.7	0.1	157.4	0.3	38.5	1.0
40	0.6	468	21.8	0.1	152.4	0.4	37.7	1.1
45	1.1	517	20.2	0.1	150.2	0.4	38.3	1.2
50	1.3	273	17.9	0.1	144.5	0.4	37.4	1.8
55	1.9	250	16.1	0.2	142.4	0.6	39.7	2.4
60	2.2	104	-	-	-	-	-	-
65	2.5	55	-	-	-	-	-	-
70	3.8	46	-	-	-	-	-	-
75	3.6	53	-	-	-	-	-	-
80	4.4	35	-	-	-	-	-	-

$L_{ADN, Poly(dA)} = 2060 \text{ bp}$								
$T$ °C	$\Delta T$ °C	Nbr Traj.	$\tau_{  }$ ms	$\Delta\tau_{  }$ ms	$R_{Exp  }$ nm	$\Delta R_{Exp  }$ nm	$L_p$ nm	$\Delta L_p$ nm
15	0.9	611	57.1	0.2	229.4	0.2	41.6	0.7
20	0.5	608	49.9	0.2	223.4	0.4	39.8	0.9
25	0.1	609	45.7	0.2	221.9	0.4	40.6	0.9
30	< 0.01	618	40.7	0.2	217.4	0.5	39.8	1.0
35	0.3	567	36.5	0.2	214.6	0.5	40.1	1.1
40	0.6	528	33.3	0.2	210.9	0.5	39.7	1.0
45	1.1	463	29.8	0.2	205.8	0.6	38.9	1.3
50	1.3	487	26.8	0.2	198.9	0.7	36.7	1.5
55	1.9	409	24.6	0.2	194.7	0.9	36.9	1.8
60	2.2	305	22.1	0.2	190.6	0.9	36.3	1.9
65	2.5	147	19.7	0.2	185.7	1.5	36.8	3.2
70	3.8	100	17.4	0.3	172.5	1.5	32.7	3.0
75	3.6	73	16.5	0.5	170.4	2.1	32.6	4.2
80	4.4	7	14.9	0.7	161.0	4.2	30.5	8.4





**Titre** Étude à l'échelle de la molécule unique des changements conformationnels de la molécule d'ADN.

*Influence de la présence de défauts locaux présents sur l'ADN et de paramètres physico-chimiques de la solution environnante.*

**Résumé** La molécule d'ADN, communément connue comme une double hélice, est un biopolymère possédant une grande variabilité structurale qui influe sur de nombreux processus biologiques. Certains d'entre eux peuvent être associés à une courbure locale ou à l'ouverture locale de la double hélice qui forme ainsi des "bulles" de dénaturation. Des changements de température ou des variations de la concentrations en sel de la solution peuvent aussi modifier globalement les propriétés physique de l'ADN.

Dans ce travail, nous proposons de combiner approches théoriques et expérimentales, ce qui correspond à mener de front expériences, simulations numériques et modèles physique, afin d'accéder à la mesure d'une courbure locale, ou à la variation du paramètre caractérisant la rigidité la molécule d'ADN sur des échantillons d'ADN présentant des défauts locaux ou soumis à diverses conditions physico-chimiques

**Mots-clés** Molécule unique, Polymère, ADN, "Tethered Particle Motion", Longueur de persistance, Simulation Monte-Carlo, Physique statistique

**Title** Study at the single molecule level of conformational changes of the DNA molecule.

*Impact of local defects included in the DNA molecule and of physicochemical conditions set by the solvent*

**Abstract** The DNA molecule commonly well-known as the double helix, is a biopolymer with a high degree of conformational and structural variability that participates in many biological processes. Some of them can be associated with local bending or local opening of the double helix that induces the formation of local denaturation "bubbles". Change in temperature or salt concentration of the solution may also globally modify the physical properties of DNA molecules.

In the present work, we propose a joint theory-experiment high-throughput approach, corresponding to manage in parallel experimental technics, numerical simulations and, physics models, in order to get access to the value of a local bend angle, or to the magnitude of variations in the rigidity parameter of the DNA molecule on DNA samples with local defects or submitted to various physicochemical conditions.

**Keywords** Single molecules, Polymers, DNA, Tethered Particle Motion, Persistence length, Monte-Carlo simulation, Statistical mechanics